

Universities' last melting pot?

The British government has at last devised a plan for the future of its institutions of higher education. That is progress, at least some of it in the right direction.

THE British government's intended policy on higher education, published last week (see page 531), is at least what it purports to be — a policy. In that sense, it should bring a welcome end to at least one of the debilitating influences of the past decade — uncertainty, much of it deliberately intended. Even dissenters will now know better where they stand. The same government's most recent previous policy statement on the same subject, Sir Keith Joseph's green paper published rather less than a year ago, does not even rate a mention in last week's white paper, which is meant as the basis for legislation yet to be devised. There could hardly be a clearer sign that ministers are anxious to wash their hands of the divisive and even disastrous recent past. The decision now to plan for an increase in the numbers of young people staying on in higher education after secondary school is also a welcome retreat by the government from the needless bickering with the universities about the proper size of the student population at a time of demographic contraction. But these welcome signs do not imply that the future of British higher education, or of the universities in particular, is one of assured contentment and stability.

The new statement is not so much a skeleton yet to be clothed in flesh as a set of declared intentions whose practicability must be a matter for conjecture. The arrangements proposed for the polytechnic sector of British higher education, more definite than those for the university sector though they may be, are a good illustration. The plan is to transfer control of these institutions from the local authorities which at present are formally in charge of them (but central government pays the bills) to a central council that will distribute such funds as are available among the separate institutions. That step, merely the extrapolation of administrative innovations in the past two years, will have to run the gauntlet of opposing local authority interests when it eventually reaches the House of Commons as legislation, most probably only after the impending general election. The government should be prepared to face out that opposition on the grounds that its proposals offer the best hope of eventually bridging the absurd gulf between the polytechnics and the universities which, in quality, form overlapping spectra.

Polytechnics

Whenever the time comes, the government will find it much harder to defend its plan that financial support for polytechnics should in future be determined contractually between individual institutions and the central council, which will have powers to "contract . . . for the provision of higher education including, where appropriate, the development of new courses and selective initiatives and research activities". The obvious result of this prescription is to give the central council detailed controls of its pensioners' academic functions. But on what basis will these educational contracts be written? Will polytechnics be offered contracts to educate people in the fields in which they seek to study, or will the contracts be tied to discipline? Following the analogy of commercial practice, will colleges be invited to tender for contracts, with the prize going to the lowest bidder? What measure of stability will there be in the system? To what extent will peer-review and other external criteria determine the

central council's willingness to back research proposals? All these questions are left unanswered in last week's white paper.

The government obviously has a soft spot for the polytechnics, which have been educating more and more young people as the years have passed, and whose costs are less than those of the universities chiefly because of the research element in the annual cost of keeping the universities in being. Can it be wise to let their future hang on mechanisms for distributing funds that have not yet been tried and tested? And, in passing, can it be seemly to set them such a poor example in the use of language by calling the central council the Polytechnics and Colleges Funding Council, thus giving further support to the reprehensible belief that any noun can be turned into a verb and that any participle can be used as a noun? Mr Kenneth Baker, the Secretary of State for Education and Science, who is something of an expert on late-Victorian literature, may have second thoughts on that decision as well.

Universities

The intended administrative framework for the universities is even less specific. The University Grants Committee (UGC) is to be replaced by a new body (called the University Funding Council) with an independent chairman along the lines recommended by the Croham committee a few weeks ago. But the government has rejected that committee's proposal for an independent commission for overseeing policy on higher education in the large, preferring to shoulder the burden on its own. Can that be wise? And while there is no explicit reference to contractual arrangements for the provision of educational services, the new council is to have powers (which the UGC has had but has rarely used) for earmarking for specified purposes (negatively as well as positively) part of the grants it will be making to particular institutions. British academics, whose plight in the past decade is a consequence of the lack of autonomy in the institutions which employ them, will be asking whether the new recipe will make them even more subservient to administrative whim.

Another question left hanging is that of how the central decision-making body will translate its grand designs into actuality? Will the present system of committees intended to demonstrate a kind of equity as among different institutions be replaced only by the kind of fairness that a bureaucracy can deliver? Will UGC's gentle attempt to nudge institutions in the direction of a more self-conscious regard for the quality and reputation of their research, which will continue until the new arrangements can be made legal two years or so from now, then be replaced by a more draconian solution of the same problem — the sad truth that British university academics' diligence and flair as researchers is not everywhere equal nor always sufficient to ensure compliance even with the spirit of their terms of appointment?

Academics would be well advised to concentrate their fire on what matters. For the system as a whole, it is a great benefit that the participation rate in Britain is now to increase; more young people will be well educated, fewer academics will be put out to grass. That is a substantial gain on the prospect just a few weeks ago. The question of contracts for the provision of higher educa-



tion needs careful watching because so much depends on the fine print in which contractual arrangements with academic institutions are reached. The government's expectation is a measure of public accountability for public funds, which is reasonable enough; the danger is that contracts will be so specific as to made academics into automata and academic institutions into sausage machines. The best hope of heading off trouble is that institutions should be more willing than in the past to open themselves to general scrutiny.

Nothing has yet been said about research except that the government sets great store by research and especially by research that has industrial application. Two sets of considerations apply. First, no part of the British university system will profit from further concentration of funds and people in a smaller number of centres. A quicker, if more arbitrary, pattern would be preferable. But in the government's and the country's best interests, the steady drumbeat of the message that economic imperatives demand applied research should be more steadfastly named for what it is — an empty and shortsighted message. Universities are better at what British universities used to do well than at providing industry with innovation; moreover, as the past few weeks of excitement about superconductivity have shown, basic research is now often nearer to innovation than the most carefully calculated projects. □

Taxing what is free

Innovative to the last, the British government is to sell radio frequencies on the open market.

HERE is a wheeze calculated to set entrepreneurial juices running. Last week the British government invented a new and exceptionally innovative scheme — selling off chunks of the radio-spectrum to the highest bidder. The logic is impeccable. Although Maxwell's equations allow for the propagation of all suitably prepared radiation in all directions without limit, geometry implies that the intensity of a broadcast signal decreases with some power of the distance between one (on a plane) and two (in free space). Accordingly, soon after Marconi's first demonstrations that radio waves will travel, there came into being an international convention by means of which broadcasting radio-frequencies were assigned to neighbouring signatories, each tagged with an appropriate broadcast power, with the objective of making sure that mutual interference is as little as possible. Usually, governments have passed on their allocations to their dependants, national broadcasting agencies and the like, after defence departments have claimed a larger share than they can use efficiently. Now the ball game may change.

Last week, the Department of Trade and Industry published a report commissioned from the consultancy firm CSP International which has recommended that the best way of using the British allocation of radio-spectrum is to put it up for auction. The department is inviting comments on the proposal, which should be regarded more as part of its continuing battle with the Home Office (which has statutory responsibility for licensing radio users) than as a way of turning Maxwell's equations to profit. That changes along these lines would make some difference is beyond dispute; CSP International is right to say that there are many uses of radio communication not now exploited because of the administrative problems of licensing; equally, it must be plain that the sale of radio-frequencies (called deregulation) would, if unregulated, be a licence to make extra-fiscal levies on commercial television companies or even further to subdue bulk users such as the BBC (British Broadcasting Corporation). Luckily, there is a formula: transmission capability is a function of bandwidth (W) and frequency (ν), but the cost of using the facility is a function of ν , say $f(\nu)$. CSP International's next job should be to determine $f(\nu)$. Then, any judge could tell what should be a proper price. □

Is negotiation strength?

Mrs Thatcher's Moscow visit has left unanswered important questions about nuclear weapons.

MRS Margaret Thatcher, the British prime minister, is not the first traveller to the Soviet Union to have come away surprised at the pleasure engendered by even a brief stay among such clever and amusing people. Again like others, she has also evidently been struck by the poignant contradictions of the place. Mr Mikhail Gorbachev, the best of its leaders for a long time, is thus still struggling to win substantial economic product from a system in which everybody nevertheless has a job. Yet it was never on the cards that she should have been able to emulate her predecessor, the late Mr Harold Macmillan (as he then was), who returned from Moscow in 1983 with the essence of an agreement on the treaty banning tests of nuclear weapons in the atmosphere. Britain was then one of the three states to have taken the initiative in arranging that there should be negotiations. On this occasion the British government is rather anxious not to be drawn into discussions about its continued and continuing status as a nuclear power.

The situation is really rather curious. The United States and the Soviet Union are busy with bilateral negotiations at Geneva whose objective is to remove their intermediate nuclear missiles from the European theatre. Broadly speaking, this is the proposal the US president, Mr Ronald Reagan, put forward at last year's Reykjavik meeting with Gorbachev, and which echoed his own earlier proposal (in 1983) then known as the "zero-option". Broadly speaking again, this line of negotiation is generally welcomed in Western Europe, chiefly on the grounds that Europe would be a safer place if there were fewer missiles that could be used by one side or the other to demonstrate its resolve without risking its own territory. France, iconoclastic as ever in matters of defence, nevertheless has insisted since 1957 that a nation with any pride must make preparations for its own defence.

British nuclear forces are sustained first by the argument that to contribute to the general defence through the North Atlantic Treaty Organization is to win influence, second by the argument that, if things go wrong, Britain will be in the same case as France. This is the spirit in which successive British governments, have said that their nuclear forces, although capable of hitting targets identical with those at which Pershing II and cruise missiles are aimed, are essentially strategic and therefore fodder for negotiation only in the context of an improved agreement (compared with unratified SALT II) on strategic missiles. One of the odd features of last week's Moscow visit is that Gorbachev appears not to have taken the opportunity of reminding his guest of this position, while she appears to have overlooked it; newspaper reports have Thatcher insisting that British nuclear weapons are not negotiable.

Can that be the case? And if so, why the turnabout? The clue lies in the meaning attached to "if things go wrong". In the old days, the meaning attached to this phrase was that there might be some always-distant set of circumstances in which the Western alliance would fall apart. Unfortunately the past few months have given most Western European governments an altogether too vivid sense that such a happening might be the result not of deliberate policy but of a kind of administrative accident. Lack of forethought, let alone of consultation, over the supply of arms to Iran, has been a considerable worry. Firing from the hip over the proper price of computer chips is another. Nobody in his senses would equate any of this with the break-up of the Western alliance; but events have shown that it is more brittle than has been supposed. Yet at some level, in an appropriate connection, the British (and even the French) nuclear forces must be negotiable. Was Thatcher, perhaps only subliminally, playing for a separate bilateral negotiation? And did Gorbachev agree? □

Slimmer US space station, same planetary probes

- Permanent occupation by 1996
- Agreement worked out with Europeans

Washington

THE National Aeronautics and Space Administration (NASA) proposed last week a scaled-back initial deployment configuration for the manned space station. Under the new plan, the space station will be permanently occupied by 1996, as promised, but a second phase, to be approved at a later date, will be required to give it all the capabilities at first intended for it.

President Ronald Reagan, a strong advocate of the space station, approved the two-phase approach late last week.

NASA's plan is now that the deployment of the space station will be carried out by 16 shuttle flights, the first of them in 1994. At the outset, the station would consist of US laboratory and habitation modules, a US polar platform, four resource nodes and provisions for experiments outside the pressurized modules. The European Space Agency (ESA) Columbus laboratory module, the Japanese experiment module and at least part of the Canadian mobile servicing system would be part of the first phase. Man-tended operation would begin in 1995, with permanent occupancy planned for 1996.

Structurally, the space station will be reduced under the plan to a single transverse boom, without the planned double keels and upper and lower booms where additional experiments were to have been mounted.

The new configuration will also lack one free-flying platform and a servicing facility. Fifty kW of power for the new initial configuration will be provided by solar cells alone. Thirty kW are needed for station housekeeping, leaving only 20 kW for users. A solar dynamic power system providing an additional 50 kW will be delayed until phase two.

Earlier this year, NASA estimated that the fully configured space station would cost approximately \$13,000 million, plus \$1,500 million for ground support, compared with the original estimates of \$8,000 million for the entire project. For its new configuration, NASA estimates that it will need \$10,900 million, with an extra \$1,300 million for ground activities.

The new plans will not affect NASA's 1988 budget request for \$767 million for space station work, now before Congress. But NASA has asked the National Research Council to conduct a technical and cost evaluation of its space station plans. This evaluation, to be completed by 1 Sep-

tember, will be used in guiding projections for the administration's 1989 budget.

Meanwhile, NASA seems to have settled with the European Space Agency (ESA) a politically sensitive question last week by agreeing to a launch timetable for two missions both bound for Jupiter. At a meeting here on 3 April, NASA administrator James Fletcher and ESA director general Reimar Lüst agreed to launch NASA's Jupiter probe Galileo in November 1989, while the joint ESA-NASA mission to the Sun by way of Jupiter will take off in October 1990. Both missions will be launched by the space shuttle.

According to last fall's revised shuttle timetable, there was only one slot for a planetary mission during the 1989 launch window for missions to Jupiter. As Galileo is a much larger and more complex spacecraft than Ulysses, keeping it stored on the ground is a more expensive proposition. But bumping Ulysses created different problems.

Europeans have been upset for some time over NASA's decision not to proceed with a sister spacecraft to Ulysses as originally planned. Moving a US space mission ahead of Ulysses seemed politically unattractive.

But as things stand, both sides agree that the present arrangements are probably for the best. When NASA cancelled development of the Shuttle-Centaur upper stage, Galileo lost the potential for a direct flight to Jupiter. When it is launched in 1989, it will fly first to Venus, then back to Earth, using those planets to change its velocity sufficiently to reach Jupiter. It will arrive in 1996.

Ulysses, on the other hand, will fly directly to Jupiter, using that planet to whip it out of the ecliptic so that it can fly over the Sun's poles. Ulysses will start returning prime solar data in 1994, before Galileo even with its later launch date. If Galileo did not fly in 1989, its next launch window would be in 1991.

Thomas Donahue, chairman of the National Academy of Sciences space science board, says that if the shuttle must be used for Galileo and Ulysses, then NASA's timetable is appropriate. But Donahue maintains that NASA is making a big mistake by not obtaining an expendable rocket as a back-up. Donahue would like to see expendable rockets for all NASA's space science missions, a direction in which NASA has shown no interest in heading.

Joseph Palca

British higher education in the 1990s

London

THE University Grants Committee (UGC), the chief source of finance for British universities, is to be reconstituted and the financially strapped research councils are to be given an extra £15 million to stave off an immediate cash crisis. Last week Mr Kenneth Baker, the Secretary of State for Education and Science, published a white paper (policy statement), *Higher Education: Meeting the Challenge* (Cm 114, HMSO, £4.50), which is meant as the basis for legislation in some future parliament.

The extra £15 million is intended to meet the 16 per cent increase of academic salaries payable from 1 December last. The Science and Engineering Research Council responded by lifting the six-month moratorium on new research grants.

The reorganization of UGC follows in outline the recommendations of the Croham committee, which on 10 February advocated a small council with a non-academic chairman. The successor body will be called the University Funding Council (UFC), while the public sector of higher education, embodied in the 26 polytechnics and larger number of colleges of higher education, will be supported by a central body, called the Polytechnics and Colleges Funding Council (PCFC).

The white paper says that one objective of the changes is "better value for money". For this and other reasons, the polytechnics will be removed from local authority control and the government plans to allow for the creation of an extra 50,000 places for full-time and part-time students in higher education.

The PCFC will negotiate contracts for the "provision of higher education", details of which are yet to be made public. There is a similar reference to "new contract arrangements" for the UFC, which says that "the government should play no part in the distribution of general funding".

The government also says that the increase of student places will mean that one in five school-leavers will have the benefit of higher education by the end of the century. These plans entail a "continuing increase of higher education spending from £3,800 million in 1987-88 to £4,100 million in 1989-90".

The government also says that the number of student places has increased by almost 160,000 (to 934,000) since 1969. The general increase of the participation rate will arise because of the expected fall in the number of 17/18-year-olds by a third in the coming decade.

Bill Johnstone

Cautious reception for British white paper on higher education

London

THE British government's white paper on higher education (see p.531) "raises more questions than it answers", according to a University of Cambridge spokesman, in a reaction echoed by universities and polytechnics throughout the country. Education administrators fear a far greater "interventionist role" in university operations and the direction of research, in the light of new proposals to tie financial support to an undefined system of contracts.

"This idea of contracts could lead to a university system directed to the needs of industry and commerce", said Sir Mark Richmond, the vice-chancellor of the University of Manchester.

The new University Funding Council (UFC) would use a system of contracting, according to the white paper, to "sharpen accountability for the use of public funds", and to "strengthen the commitment of higher education institutions to the delivery of educational services" agreed upon by the UFC.

The report gives no details about what would be included in the contracts, and educators are unclear if the new system would be linked with the number of students in a particular subject, their qualifications or other details. "Is the university going to have to produce 50 graduates, of whom ten get first class degrees", one administrator asked. "If the department doesn't deliver, will it be penalized?"

The government clearly intends to direct education and research into areas required by industry, with the underlying assumption that the increasingly diminishing resources would be selectively distributed. "But", said a University of Edinburgh spokesman, "if the emphasis is more towards applied research and education, it creates concern among the areas not directly related to immediate industrial well-being, in the arts, the social sciences, and to some extent parts of the sciences."

Under the new policy, Britain's polytechnics and colleges would be freed from the control of local education authorities, a proposal generally welcomed by the polytechnics. "Some local authority procedures are not geared to the provision of education", explained the director of the 6,000-student Coventry Polytechnic, Geoffrey Holroyde.

But the distribution of financial support to Britain's polytechnics would be determined by a new Polytechnics and Colleges Funding Council, which would, according to the white paper, "have a strong industrial and commercial element". In

addition, the Secretary of State for Education and Science would have reserve powers of direction, a stronger role than he would take in the University Funding Council.

"This move on the polytechnics is highly political", said Richmond. "The government has been uneasy about the way the local authorities spend the money." There is some concern among polytechnic administrators that the government will intervene in the direction of these institutions to provide a kind of manpower planning for industrial and commercial interests.

One aspect of the white paper that has been welcomed is the decision to increase the number of students over the next decade, although the numbers will fall again by the mid-1990s. But the projected increase of 24,000 full-time and 27,000 part-time students has been criticized by the Committee of Vice-Chancellors and Principals as "too unambitious". Lack of funding for the increased numbers is also of concern — "the white paper is totally lacking in any indication of how the extra students are to be funded", said Sir John

Burnett, vice-chancellor of the University of Edinburgh, who points out that there is no proposal to increase student maintenance grants.

The government has also proposed increased autonomy for educational institutions in Scotland, with the provision for a Scottish Committee to provide input to the UFC on Scottish needs. This move is seen as a compromise, as the Scottish Tertiary Education Advisory Council recommendation for a separate funding system has not been supported by the Scottish universities. "We'll be very interested in seeing just what will be the standing of the input from the Scottish Committee", said a University of Edinburgh spokesman.

The government's recommendations on higher education are presented in a glossy booklet full of colourful graphics, in contrast to the usual stark white paper. It is seen by some education administrators as an election manifesto, because its provisions could not be implemented before the British election.

Higher education officials agree that there are many details to be supplied about the new proposals, before there can be any move to change policies. "There is going to be some very active lobbying going on," one official concluded.

Kathy Johnston

Judge confirms injunction in sandwich assay patent suit

Washington

IN the latest round in the court battle between Monoclonal Antibodies Inc. and Hybritech Inc. over the patent for sandwich assays, Monoclonal Antibodies has suffered a severe blow. As a follow-up to the court decision that upheld the contested Hybritech patent for the assays last fall, a US judge has confirmed that a preliminary injunction will be issued against Monoclonal Antibodies Inc. The injunction will prevent Monoclonal Antibodies from selling products that comprise 80 per cent of its product sales.

The dispute between Hybritech, a subsidiary of Eli Lilly, and Monoclonal Antibodies centres on 'sandwich assays', a technique for diagnosing physical conditions such as disease or pregnancy using monoclonal antibodies. Hybritech sued Monoclonal for patent infringement in 1984, and its patent was overturned in subsequent litigation due to obviousness. Hybritech appealed, and the appellate court upheld the Hybritech patent (see *Nature* 324, 506; 1986).

The preliminary injunction against Monoclonal Antibodies mainly affects its OvuStick test for determining ovulation, and Advance, a pregnancy kit made by Monoclonal and marketed by Ortho Phar-

maceutical Corporation. The cost to Monoclonal of the injunction against sales of these products is estimated to be \$1.1 million a quarter.

Thomas A. Glaze, president of Monoclonal Antibodies, says the company will now focus on the development of products



that will not be encumbered by patent problems. He hopes to introduce tests for sexually transmitted diseases and the detection of drug abuse in late 1987 or early 1988.

But the question of damages has not yet been solved, and a large award to Hybritech could further debilitate Monoclonal Antibodies. Industry analysts have already pinpointed Monoclonal as a possible takeover target.

Carol Ezzell

Settlement on AIDS finally reached between US and Pasteur

Washington

WITH a mixture of elation and relief, the Institut Pasteur and the US Department of Health and Human Services (HHS) last week signed an agreement ending a long and at times divisive dispute over who should rightly claim credit for the discovery of the virus causing AIDS (acquired immune deficiency syndrome) and the commercial blood test that screens for antibodies to that virus.

President Ronald Reagan and Prime Minister Jacques Chirac announced the settlement in a brief ceremony at the White House on Tuesday, 31 March.

By the terms of the settlement, both sides will contribute 80 per cent of all royalties received from sales of the antibody blood test to a foundation that will

support research on AIDS and other diseases caused by a human retrovirus. The new foundation will have six trustees — three appointed by the chairman of the board of directors of Pasteur and three by the Secretary of Health and Human Services. The foundation will contribute 25 per cent of its grants for research on AIDS in “the less developed countries of the world”. Based on present estimates of worldwide royalties, the new foundation should have an income of between \$4 and \$5 million this year.

The settlement also includes a “definitive” but “by no means exhaustive” scientific history (see *Nature* 326, 435; 1987) of the contributions of both laboratories to human retrovirology and the development of an AIDS antibody test-kit. Robert Gallo of the National Cancer Institute and Luc Montagnier of the Institut Pasteur and their colleagues agree not to “make nor publish any statement which would or could be construed as contradicting or compromising the integrity of the said scientific history”. Lawyers for both sides agree that those not bound by this restriction are likely to give the history careful scrutiny.

The settlement will end several outstanding legal disputes between the two parties. Both Gallo and Montagnier had submitted patent applications for a blood test to detect antibodies to the AIDS virus. Although Montagnier's US patent was filed in December 1983 — four months before Gallo's application — the US patent office awarded Gallo a patent in May 1984 without acting on the Montagnier patent application.

The patent office last year initiated a procedure known as an interference that would have determined whether Montagnier's patent deserved priority. But by the terms of the settlement, the interference will be “terminated without an award of priority”. The blood test will be regarded as a “joint invention” between both laboratories, and the patents will be amended to show the names of both Gallo and Montagnier and their colleagues.

The settlement will also end a dispute over a sample of the AIDS virus isolated by Montagnier and given to Gallo's laboratory in September 1983. In accepting the sample, Gallo and his colleagues promised to use it for investigation only.

Institut Pasteur had claimed that Gallo and colleagues used this isolate in developing their commercial blood test. Although a US claims court dismissed the Pasteur case last July, Pasteur's lawyers appealed, and on 9 March this year an appeals court reversed the initial dismissal and ordered the lower court to reconsider

the case. That case will now be dropped.

The Institut Pasteur will also agree to drop a request under the Freedom of Information Act to receive documents relating to the patent suit and US licensing arrangements.

Both HHS and Pasteur will retain their current licensing arrangements, but in future each side will be free to license the other's technology. This includes not only the original antibody test-kit patent, but also a patent awarded to Gallo this year for growing the AIDS virus in permanent cell lines. The agreement also states that “improvement technology” — technology that would improve on the basic test-kit technology, possibly using recombinant DNA technology — developed by either side in the dispute will be equally available to both sides' licensees.

Although some on both sides of the dispute still believe that a court battle would have vindicated their position, the settlement was regarded with jubilation by the principal scientists involved. In Gallo's laboratory in Bethesda, they broke out Korbel champagne and Montagnier, travelling in Australia, said that he would be toasting them in spirit.

Lawyers for both HHS and Pasteur agree that the decision late last year to designate a single primary negotiator for each side cleared the way to reaching a final settlement. Ira Millstein of Weil, Gotshal & Manges represented Pasteur's interests, and HHS general counsel Ronald Robertson negotiated for the US side. Both sides agreed that Jonas Salk played a crucial role in bringing the settlement negotiations to a successful conclusion.

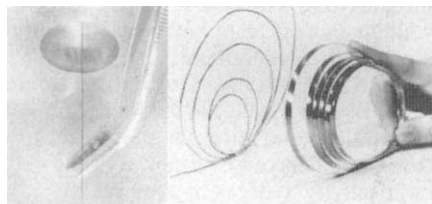
Joseph Palca

Superconducting ceramics race on

THE race to turn the new superconducting ceramics into industrial products may be only a couple of months old but the first successes are already being scored.

On the left, 0.2-mm diameter wire from the Argonne National Laboratory in Illinois, United States; on the right, 0.6-mm wire and 5-mm wide, 0.1-mm thick tape from Toshiba Corporation in Japan. All are made from yttrium-barium-copper oxide, which superconducts at liquid nitrogen temperature.

The wire has potential in high-field



magnets, but is handicapped by the low currents at which superconductivity ceases. Toshiba's sample fails at 6 A cm^{-2} and Argonne's at around 5 A cm^{-2} — several orders of magnitude too low for practical application.

The wires above are brittle and have to go through an expensive forming process before they are fired and made superconducting. But researchers at Argonne have shown that if thin enough, the fired wire can be bent.

Almost simultaneously with Toshiba's announcement, Matsushita Electric Industrial Company's Central Research Laboratory claimed to have developed a superconducting monocrystal ceramic film of lanthanum-strontium-copper which may have application in large-scale integrated circuits. □

German AIDS registration row

Munich

A CONFERENCE of the health ministers of the eleven West German *Länder* (states) has rejected Bavaria's call for mandatory registration of AIDS (acquired immune deficiency syndrome) carriers. It is therefore very unlikely that Bavaria will succeed in incorporating its far-reaching proposals for AIDS registration and testing into West German law.

The Bavarian government nevertheless plans to press ahead with its programme, which includes mandatory testing for members of high-risk groups such as prostitutes and intravenous drug users, but also for foreigners intending to live in Bavaria. These plans have led to a migration of people infected with AIDS into the neighbouring state of Baden-Württemberg and Hesse. One report indicates that as many as 20 per cent of those seeking treatment and counselling in Ulm, a city just outside the Bavarian border, come from Bavaria.

Steven Dickman

Fast-breeder test reactor faces licence refusal in Germany

Munich

THE West German state of North Rhine-Westphalia intends to refuse the "overall approval" necessary for operating the nearly completed fast-breeder reactor at Kalkar. This was announced in Düsseldorf on 1 April by the state Economics Minister Reimut Jochimsen, a Social Democrat (SPD).

The 300-MW plant at Kalkar is 70 km west of Düsseldorf, not far from the Dutch border. Under construction since 1972, it has cost roughly DM6,500 million so far. The bulk of the funds has come from the Ministry for Research and Technology (BMFT). The reactor is meant as a test reactor for fast-breeder technology.

What Jochimsen says does not necessarily mean the end for the fast breeder. He has given the plant's principal owner, Schnellbrüter-Kernkraftwerksgesellschaft (SBK), until 30 June to respond to his complaints. Jochimsen claimed that there have been 2,000 unauthorized alterations in the reactor's construction and that Kalkar's design is uncomfortably similar to that of Reactor 4 at Chernobyl.

The plan to prevent Kalkar from going into operation was criticized as a political ploy. SBK's technical director Werner Koop pointed out that the 2,000 changes cited have occurred over a period of 15 years, during which time some 500 tonnes of reports have been delivered to the state government. Most of the changes, said Koop, do not even require approval from

the state licensing authorities.

Rolf Hüper, a scientist in West Germany's Nuclear Research Centre (KfK) at Karlsruhe, said that all of the comparisons of Kalkar with Chernobyl have been dealt with. For example, it is not so alarming that the plant at Kalkar uses sodium as a coolant in the light of experience at the



The nearly completed fast-breeder reactor in Kalkar.

Soviet BN 350 fast breeder on the Caspian Sea, where a sodium-water explosion "in tonne quantities" occurred without any release of radioactivity.

Elections are not due in North Rhine-Westphalia until 1990, but Social Democrats were facing a tough election in the neighbouring state of Hesse on 5 April. Jochimsen's announcement may have been a way to win support for the SPD from anti-nuclear voters, particularly supporters of the Green party.

A decision to refuse to license the Kalkar plant would put North Rhine-Westphalia's government on a collision course with the West German government. But

the federal Environment Minister, Walter Wallmann (Christian Democrat), has not yet said that he would try to force Jochimsen to license the plant, saying instead that he would wait until the summer before deciding on a course of action. If Jochimsen stands firm, the dispute could wind up before the German Supreme Court at Karlsruhe, just like the suit brought by the state of Hesse challenging West Germany's "plutonium economy".

BMFT would like to see the reactor licensed. A spokeswoman for the ministry said, "we have an interest... in seeing the breeder go into operation, but that does not mean we will finance it".

SBK's Koop said that the current financing of the Kalkar reactor will hold out until April 1988. Each month's delay, however, costs an extra DM10 million.

In the end, costs rather than politics may decide the fate of the reactor. With the cost of both Kalkar and the reprocessing plant at Wackersdorf far above the original estimates, the West German government may have to rethink its plans to build a larger breeder in partnership with France.

Steven Dickman

The costs of nuclear decommissioning

London

THE means of determining the cost of decommissioning a Magnox nuclear power station and the period over which that cost can be spread has been seriously challenged by the UK House of Commons Energy Committee. In a report published on 1 April the committee concludes that the only real way to assess the economic viability of a nuclear plant is by "discounting of future decommissioning costs".

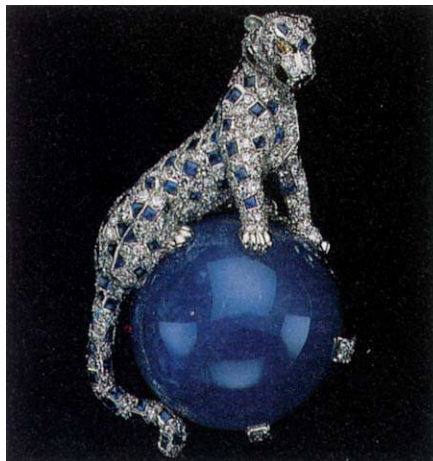
Central Electricity Generating Board (CEGB) policy calls for the final stages of dismantling and disposal of the reactor core to be delayed by a century, allowing radioactivity to decay to a low level.

The committee concludes that there is no realistic alternative in assessing the economics of a reactor but it has challenged the relevance of this method over a period longer than 50 years. And says that "it is essentially an accounting exercise which pays no regard to questions of moral responsibility to future generations. It is scarcely possible to believe that the coal owners of 1887 made assumptions about the value to us of coal."

In evidence to the committee, the CEGB had challenged press reports which claimed that decommissioning would cost £2,700 million (1984-85 figures) a station, they estimated a maximum of £330 million and environmental lobby groups provided estimates of between £500 million and £1,000 million based on 1986 figures. Variations are also due to alternative ways of performing the decommissioning.

Bill Johnstone

Unexpected windfall for Pasteur Institute



London

LAST week's auction of the Duchess of Windsor's jewellery has brought a windfall for the Pasteur Institute in its centenary year. In her will, the duchess directed that almost all the proceeds of the sale, expected to be about £5 million, should go



to the Pasteur. The auction raised £31 million. A new £8-million department for the study of retroviruses and AIDS (and, perhaps, cancer) will go ahead. The rest of the money may simply be invested to support the department, which will employ 50-100 scientists.

Peter Newmark

Healthy scientific environment promoted by society in India

New Delhi

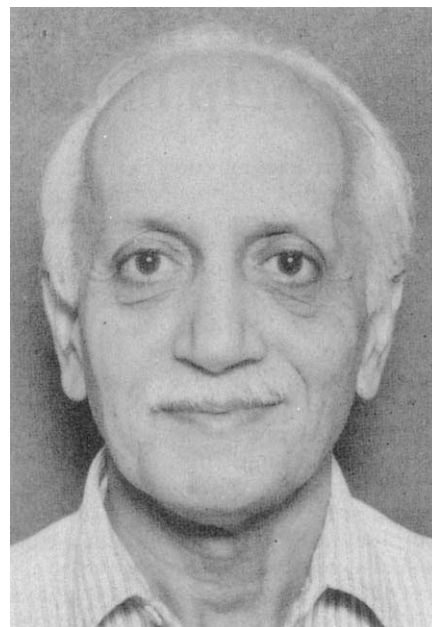
SOME 107 prominent Indian scientists have formed the Society for Scientific Values (SSV) to promote integrity, objectivity and ethical values in the pursuit of science. Its chief aim is to evolve a healthy scientific environment "free from prejudices, bureaucratic formalisms, dishonesty, propagation of unsubstantiated research claims, suppression of dissent, showmanship, sycophancy and political manipulation".

The architect of this new movement is Dr Avtar Singh Paintal, director-general of the Indian Council of Medical Research (ICMR). Paintal says that the climate for scientific research in India has deteriorated since independence not because of inadequate research facilities or salaries for scientists but due to the lack of a healthy scientific environment. And the

society's chief objective is to promote, by personal and collective effort, the ethics and norms of science "not only for the progress of science and technology in the country but also for national character".

The membership of SSV is open only to scientists with a clean record who are serious in promoting the society's ideals in the organizations in which they work. Members should never have plagiarized, made false claims or indulged in any kind of dishonest activity and should not have allowed their names to be included in publications in which they have not been actively involved.

The movement is gradually showing results, says Paintal. The eligibility criterion for authorship of papers is becoming widely accepted. The society has also persuaded the government to set up conference facilities for scientists: SSV prohibits



Dr Avtar Singh Paintal.

its members from attending conferences in expensive hotels beyond the reach of students and university scholars.

The practice of giving prizes year after year to the same group of politically influential scientists, who have long since ceased to do research, has become a scandal in India, and Paintal says that SSV hopes to put an end to this unhealthy tradition.

Paintal, has also redefined the goals of his own organization, ICMR. Although ICMR will continue to finance advanced research in medicine, it will chiefly be concerned with the application of existing technologies in solving India's health problems. Paintal's two priorities are the control of mosquitoes using the environmental approach and family planning using contraceptive pills.

According to Paintal, India has made a mistake by not promoting the pill. Fewer than 300,000 women in the whole of India at present use this method. At the instance of ICMR, the government is about to launch a major campaign advocating the use of the pill by women and of condoms by men.

On the control of malaria and filariasis, Paintal sees no need for new research. He says the only way to control these diseases is through elimination of the vector by environmental control, filling up ditches and improving sanitation. This approach has helped drastically to reduce the incidence of filariasis in Pondicherry and malaria in the Khera district of Gujarat. For the past 30 years, India has spent nearly half its health budget in the war against mosquitoes but without success. Paintal says that ICMR has shown that the battle can be won by simply cleaning up the environment instead of spending money on research on drugs, vaccines or insecticides.

K. S. Jayaraman

Another large telescope to be sited on Hawaii's Mauna Kea?

Washington

THE National Optical Astronomy Observatories (NOAO) have chosen Mauna Kea in Hawaii as the prospective site for a huge new instrument consisting of four harnessed 8-m telescopes. The summit of this extinct volcano already houses one of the world's largest family of telescopes.

The National New Technology Telescope (NNTT) is the most ambitious of several current proposals espoused by various sections of the US astronomical community. Its four mirrors, each much larger than anything cast or ground so far, can be made to function as four separate telescopes, or as a single telescope or as an interferometer. They will be made by a technique called spin-casting, in which molten ceramic is rotated at a constant speed as it cools, so that the solidified mass takes on a parabolic curve. The blank is then figured by traditional grinding and polishing.

A 2-m mirror has been made by this method, and preparations for making a 3.5-m blank are almost complete. If successfully cast and polished, the 3.5-m mirror will be used by the five-university Astrophysical Research Consortium for a new telescope at Apache Point, in the Sacramento Mountains of New Mexico.

The rival to Mauna Kea was Mt Graham, about 80 miles north-east of Tucson, Arizona. A comprehensive study was made of the scientific virtues of both sites, and Hawaii was preferred, according to NOAO, solely because its superior atmospheric conditions permit higher

quality astronomical observations. But there was opposition to the Arizona site by environmental groups, particularly the Sierra Club. Mt Graham is remote and unspoiled, and has its own unique species of squirrel. Remoteness is, of course, a benefit for astronomy; observations at Kitt Peak are beginning to be marred by diffuse artificial lighting from Tucson and Phoenix.

The NNTT remains an astronomer's dream, with an expected cost of about \$150 million. Present National Science Foundation (NSF) spending, at a rate of \$2 million a year, supports only technological development, of which a good deal is needed. As well as using mirrors of unprecedented size, the NNTT will demand control systems and a mounting of largely untested abilities. Feasibility studies of these problems, says Pat Bautz at NSF, are aimed at producing a formal construction proposal within a few years. NSF will then decide whether to allocate funds from its regular astronomy budget.

These doubts are behind a feeling that it might be better to keep the NNTT and other big projects in abeyance for a few years while the privately funded Keck telescope gets under way. Construction of this device, a 36-element multiple-mirror telescope, will itself explore many of the technical problems the NNTT will have to face. The Keck project should be finished in 1992, and there is an understandable reluctance to commit federal funds before the verdict is in on that venture.

David Lindley

US experimental drug rule change may help biotechnology

Washington

DRUG companies may now be able to sell, rather than provide free of charge, experimental drugs designed to treat serious or life-threatening diseases in advance of US Food and Drug Administration (FDA) approval, if proposed rule changes by FDA are adopted. This is one of the most controversial changes to the Investigational New Drug (IND) regulations, and could provide important benefits to small drug manufacturers.

The rule changes are the result of an investigation by the Presidential Task Force on Regulatory Relief, whose chief purpose was to give pharmaceutical companies increased incentives to develop and test new drugs for the treatment of serious diseases, particularly AIDS (acquired immune deficiency syndrome). Freedom to recoup a portion of drug development costs during clinical trials may be an important incentive: the cost of developing a new drug and taking it through clinical trials is estimated at \$100 million.

The proposed changes would permit a pharmaceutical company to charge for drugs administered to patients participating in clinical trials and to others eligible. But the drugs must have passed at least Phase I clinical trials, where safety to humans and dose amount is evaluated, and probably Phase II trials, where drug efficacy is tested. Drug companies would be prohibited from actively marketing the drug and would have to continue to work toward final FDA approval.

Only drugs being developed for the treatment of serious or life-threatening diseases would qualify under the proposed new regulations. Diseases such as Alzheimer's and multiple sclerosis would fall in the "serious" category, while AIDS, and certain uncontrollable cardiac arrhythmias, would be "life-threatening".

Before selling an experimental drug to patients in a clinical trial, a drug company would also have to convince FDA that the sale is needed for the sponsor to undertake or continue the trial. Companies would be able to sell drugs to patients not participating in clinical trials without restriction as long as the agreed protocol was followed.

Although the proposed regulation would benefit affected patients by stimulating innovations in drug development, it may also impose a financial hardship on them. Most health insurance companies will not pay for drug therapy not approved by the FDA. But a spokesman for the Health Insurance Association of America pointed out that if no other treatment is available, some companies might pay on a case-by-case basis. FDA has said that it

will not allow excessively high prices for experimental drugs, and will withdraw its approval if the price for a drug is unfair.

The new ruling would have a very favourable effect on the biotechnology industry, and particularly on small companies, according to Richard Godown, president of the Industrial Biotechnology Association. The development of drugs produced by genetic engineering is especially costly. Several smaller companies have drugs in clinical trials that may qualify under the proposed regulations.

According to Jim McCamant, editor of the *Medical Technology Stock Letter*, anti-cancer drugs are likely candidates. Cetus has its tumour necrosis factor in Phase II trials, Ribic Immunochem has adjuvants for the treatment of a several cancers in Phase I and II trials and Xoma has immunotoxins to treat melanoma in Phase II tests. Genentech's tissue plasminogen activator (tPA) might also qualify, although it is expected to win FDA marketing approval soon.

Because of the controversy surrounding the proposed changes, FDA is accepting public comments until 20 April. If approved, companies could begin selling experimental drugs as early as 17 June this year.

Carol Ezzell

OTA gives Congress options on cell-line donor ownership

Washington

THE Office of Technology Assessment (OTA), the technology policy analysis arm of the US Congress, last week issued a report* intended to help Congress to clarify a now-fuzzy area of biotechnology — the ownership of human tissues and cells. The contents of the report could lead to new regulations and procedures for the collection and use of human tissues in research and product development.

There is at present no legislation expressly governing human cell and tissue culture. Although this has not been considered a problem in the past, rapid developments in biotechnology in the past several years using human products have caused concern. At issue is whether the original human donor should share in the commercial benefits to be reaped from techniques employing human cells, and the legal and ethical questions arising from selling human tissues. The technologies concerned are the development and maintenance of human cell lines, hybridomas for making monoclonal antibodies and recombinant DNA techniques such as gene probes.

Under the present system, the benefits of commercial products developed from human cell cultures are not shared with the person who donated the cells, who is usually not aware that his tissues have been put to commercial use. The tissues used in both basic research and the development of products from human cells are likely to be cast-offs from surgery. Clonetics Corporation, for example, uses human skin from elective plastic surgery as a source for its normal human epidermal cell cultures.

But the growth and manipulation of human cells is often described as more art than science. Human cells are fastidious in culture, and attempts to create immortal cell lines from non-tumour tissue are

rarely successful on more than 1 per cent of occasions with some tissue types. In the case of sophisticated technology using human cells, such as the creation of hybridomas, it is the efforts of the researcher that make a profitable product.

Given these considerations, the question arises whether the human donor is entitled to a share in the profits. Industry researchers claim that it would be extremely tedious, and sometimes next to impossible, to trace the development of a commercial product back to its human source for the purpose of profit-sharing. Tissues and cell cultures are now freely exchanged between researchers.

The OTA report describes several options for Congress to explore in resolving these issues. If the ethical rules that determine compensation for human-tissue donations are deemed appropriate, Congress could enact procedures to force researchers to keep detailed records on the sources of tissues and cells. Donors could be reimbursed for the 'service' of donation, or could receive no compensation at all unless their donation resulted in the development of a profitable product. Conversely, Congress could rule that the ban on the sale of human organs by the National Organ Transplant Act extends to human tissues and cells, thus barring patients from recompense for tissue and cell donations.

Representative Robert A. Roe (Democrat, New Jersey), chairman of the Committee on Science, Space and Technology, has asked the Department of Health and Human Services to evaluate the options outlined by the OTA report regarding informing patients about the uses of their donated tissues.

Carol Ezzell

* *New Developments in Biotechnology: Ownership of Human Tissues and Cells* is available as OTA-BA-337 from the US Government Printing Office, Washington, DC.

UK release of four genetically manipulated organisms planned

London

POTATOES are likely to be the first genetically engineered plants to be grown in field trials in Britain. Two proposals for planting in May or June have been made to the UK watchdog on deliberate release experiments, the Advisory Committee on Genetic Manipulation (ACGM). But both have run into some problems which will need to be resolved by the next ACGM meeting, in mid-May, for optimal plantings this year. The same meeting may also be the last chance for two other releases, one of bacteria and the other of viruses, to go ahead this year.

Only two of the proposals involve organisms that have been produced by genetic engineering in the sense of using recombinant DNA technology to alter their genetic make-up. One is a potato experiment proposed by the Plant Breeding Institute in Cambridge; the other is the second phase of a project planned by the Institute of Virology in Oxford. The organisms involved in the other two experiments have been produced at Rothamsted Experimental Station by more classical means of genetic manipulation which have nonetheless been deemed to fall within the purview of ACGM, somewhat to the initial surprise of the investigators.

The ultimate aim of the Cambridge work is to produce disease resistance and, perhaps, increased protein in potatoes, but the initial trial is only of a prototype containing convenient, rather than useful, genes. One gene confers antibiotic resistance on the plant cells, which has allowed their selection in the laboratory. The other gene encodes a bacterial enzyme that can easily be assayed.

Plants with both genes have been obtained by tissue culture and tested in glass-houses. The plan now is to grow about 2,000 of them in a small plot this year and then to replant the tubers next year, which will provide a better measure of any effects of the genes on the yield and size of the potatoes and the gross morphology of the plants.

Details of the proposed planting were sent last December to ACGM, which responded only in March with several suggestions and an amended plan will be reconsidered in May. A variety of precautions are included to prevent any spread of the plants beyond the experimental plot.

Similar precautions have had to be incorporated in the Rothamsted potato experiment even though the plants are the result only of fusions between cells of two species of potato, just one of which is resistant to the leaf roll virus. As hoped, some of the hybrids are also resistant.

The other Rothamsted proposal involves *Rhizobium*, the nitrogen-fixing bacterium of leguminous plants. The experiment is financed by a European Economic Community programme for risk assessment in biotechnology and has both French and German collaborators. To test the extent of genetic transfer between rhizobial species in the soil a *Rhizobium* has been genetically marked with a DNA transposon from a different bacterial genus and as one of ACGM's concerns is to prevent this kind of genetic transfer, the experiment was bound to be problematic. The original plan is being redrafted in response to considerable, if delayed, criticism by ACGM. Ironically, similar experiments were carried out by a member of ACGM before it came into existence.

The Institute of Virology's experiment is the second phase of a project that started last year with what is, so far, the only deliberate release of genetically engineered organisms in Britain. Last summer, institute scientists released a baculovirus that was genetically marked so that its survival could be monitored. Ultimately the plan is to alter the virus to improve its value as a biological pesticide. Such viruses will also be altered so as to limit their survival and it is one such 'suicide' virus that may be tested this summer.

Although ACGM is a voluntary body that gives advice rather than approval for experiments, none of the experimenters will go ahead this year if the committee is not satisfied with their modified proposals. As ACGM is building up guidelines on a case-by-case basis, delays are not surprising but given the slowness with which ACGM has reacted to most of the outstanding proposals, there will still be a good deal of frustration if any are delayed by a year.

Peter Newmark

Japan faces both ways on timber conservation in tropical forests

Tokyo

THE International Tropical Timber Organization (ITTO), a world body mandated to promote timber trade and preserve tropical forests, began its first council meeting in Yokohama in a state of financial crisis. By the end of the five-day meeting many member countries, including the United States, had failed to pay the dues essential for the secretariat's operations. But Japan, the biggest financier of destruction of the world's tropical forests, came away with a glowing image having pledged by far the largest voluntary contribution to the organization's special project fund.

ITTO was established last summer by all the major tropical timber producing and consuming nations under the International Tropical Timber Trade Agreement, which includes among its principal goals the preservation of tropical forests (see *Nature* 312, 493; 1986). Japan lobbied hard to host the organization's headquarters, and after a year-long squabble, Yokohama was chosen over Amsterdam, Jakarta and Rio de Janeiro.

Some environmentalists expressed concern at the choice of Japan, fearing the host might compromise the organization's conservation objectives: Japan gobbles up nearly half the world's tropical timber trade, much of the imports being wasted on disposable chopsticks (11,000 million pairs a year) and panelling for concrete, and vast tracks of forests in South-East Asia have been laid to waste to satisfy Japan's appetite for tropical hardwood.

But at the council meeting, Japan was

the first to step forward with a handsome contribution of \$2 million to finance projects on reforestation, sustainable management and economic/market information gathering. Other nations were slow to respond to the Japanese move. And so a group of environmental organizations, including Friends of the Earth and the World Wildlife Fund, chipped in more than \$12,000 to try and shame others into action. Switzerland quickly came up with \$1 million, followed by the Netherlands with \$600,000. And at the end of the meeting, \$640,000 was allotted to establish in-house statistical capability, and to do preliminary work on ten potential projects before the council's next meeting in the autumn.

But the projects may never get off the ground because several member nations have failed to pay the membership dues essential for everyday running of the organization. By the end of the meeting on 27 March, nearly \$700,000 was still owed, \$50,000 by the United States.

Charles Secrett of Friends of the Earth International, an official observer at the meeting, while welcoming the voluntary contributions from Japan, Switzerland and the Netherlands, condemned the non-paying member nations. "It is beyond belief that the whole work programme of the ITTO may fail simply because a few selfish countries refuse to pay money they owe. By miserly skimping on the small change, the defaulters are threatening the future of the one international organization that could save the world's tropical forests."

David Swinbanks

Pursuit of relevance in research

SIR—In your leading article "How to get back from the dead" (*Nature* 325, 561; 1987), with whose general argument I largely agree, you say, of the reorganized University Grants Committee (UGC) proposed by the Croham Committee, that an "obvious danger is that such a council might shoot off in an eccentric direction, the pursuit of what is fashionably called 'relevance', not excellence, for example".

If by 'relevance' you mean that which the private sector will pay for and/or is commercially attractive, then I fear that you perpetuate one of the most destructive myths ever to infect British universities, namely that the excellent must be irrelevant to commercial and industrial interests and that anything that the universities do in support of those wider interests cannot be excellent in the academic sense. I believe that excellence and relevance often go hand-in-hand and indeed can depend on one another. Thus, in the past five years, the income (in real terms) that this university has received from self-financing courses and conferences (by definition 'relevant') has increased by 973 per cent; the income from our associated companies (Salford University Business Services Ltd and Salford University Civil Engineering Ltd and very relevant to us at least) has increased by 580 per cent and our income from research council grants by 95 per cent. As during those five years it is generally recognized to have become more difficult to obtain funding from the research councils, I find it difficult to believe that our pursuit of 'relevance' has been at the expense of research 'excellence'.

It has long been accepted, of course, that degree programmes of the kind that this university offers whereby an academic education is combined with professional training are best taught by those with some knowledge of the context in which those skills are to be practised. But closer examination of our accounts for last year will reveal that our departments have spent, in aggregate, £900,000 from the monies they earned (by definition on 'relevant' work) on their teaching activities and I have just sanctioned the expenditure of £150,000 by our Teaching Committee from the profits covenanted to us by Salford University Business Services Ltd, in order to support excellence in our teaching activities. Thus, not only is research excellence at this university not incompatible with 'relevance', teaching excellence — at least as measured by access to up-to-date equipment and more favourable staff-student ratios — is becoming dependent on 'relevant' activities too. This, by the way, is one very good reason why the UGC as well as the research councils should continue to support research in

the universities because, as we have shown, the profits from the 'relevant' work that stems from our research is increasingly needed to maintain excellence in the teaching activities that are undoubtedly the UGC's concern.

JOHN ASHWORTH
(Vice-Chancellor)

University of Salford,
Salford M5 4WT, UK

Plight of postdocs

SIR—Dennis Murphy in his letter on the plight of British postdocs (*Nature* 325, 478; 1987) raises some pertinent points. The sentiments he expresses, however, bear the hallmark of one safe in the knowledge that he has security of tenure. The problem of salary grading and structure for research staff is clearly one that needs to be addressed, but removing the anomalous age-related scheme to replace it with a flat-rate scale is a recipe for disaster. For if the age-related salary structure is to be removed for fixed-term contract staff, ought it not be removed for tenured staff also?

Murphy obviously sees the postdoc position as a vocational one. How else could he believe that a postdoc in his 40s would be happy to be employed on the salary of his junior counterpart fresh from a PhD? He notes that one of the reasons for the 'brain drain' is "the greater scientific and pecuniary opportunities that exist abroad". Why, therefore, does he assume that a mature postdoc in Britain would be any happier to work for a junior postdoc's salary simply because it's a job?

Postdocs soon discover the difficulties of obtaining mortgages with fixed-term contracts and the prospect of an uncertain nomadic existence does not provide the stability needed to start a family. Unless we are to deny our fixed-term contract staff the same hopes and aspirations held by contemporaries in industry and tenured positions, the removal of the 'safety net' of age-related salaries is not a viable proposition.

A. PARSONS

Department of Biology,
University of York,
York YO1 5DD, UK

SIR—I wonder whether Denis Murphy (*Nature* 325, 478; 1986), who writes about postdoctoral salaries, would be so content if it were suggested that the annual increments enjoyed by tenured lecturers should also be abolished, and that all lecturers, of whatever age and experience, should be paid the same salary.

The whole point of the age-related scale is that it is supposed to reflect the usual correlation of age with practical labora-

tory experience. The more experienced a postdoctoral worker, the more useful he or she is to a research team. Industry recognizes this, and pays people more when they have more useful experience. In the United States, the 'tenure-track' system allows researchers to progress through at least a semblance of a career structure as they gain age and experience. In Britain, because there is no career structure for scientific researchers, there is no way for highly qualified and experienced personnel to remain in the profession. High-quality researchers will not be persuaded to stay by the prospect of continued short-term employment at the lowest possible salaries. What we now have is university research 'at the bench' by newly-qualified inexperienced postdocs while a small pool of experienced personnel who have lecturerships spend all their time on teaching, administration and the hunt for new grants, and rarely, if ever, do any research. Surely there is a better way than this?

MICHAEL D. BARON

Department of Biochemistry,
University of Leicester,
Adrian Building
University Road,
Leicester LE1 7RH, UK

Confining AIDS

SIR—There is little doubt that the incidence of AIDS (acquired immune deficiency syndrome) is surpassing all expectations and can only increase still further as it reaches Asia. It has indeed reached India as well as Thailand, where a member of the US Embassy personnel was infected.

Special interest groups, such as homosexual associations, are highly organized in purchasing the influence of legislators whose duty it is to protect the public in general. Civil liberties unions, too, are on the side of the spreader of AIDS.

The decision of the Indian Council for Medical Research (see *Nature* 324, 294; 1986) to institute blood tests for foreign students is a good beginning in the attempt to limit the spread of AIDS. But the main source of infection must obviously be sought in tourists. A visa should be given only to AIDS-negative applicants in all countries where AIDS is not yet widespread. This would be more easily accomplished in the countries, usually the United States and Europe, where most of the tourists originate and where facilities would permit the no-AIDS certificate to be part of the visa application procedure.

M.K. AGARWAL

Laboratoire de Physio-Hormono-
Réceptérologie,
Université Pierre et Marie Curie,
15 Rue de l'École de Médecine,
75270 Paris Cedex 06,
France

Environment and foreign aid

The well-known jibe that environmental protection is a self-indulgence of the rich is particularly apposite on the international plane. But the corollary is more foreign aid.

CURIOUS transformations are being effected in the environmental movement. That is one reading of the elections last weekend in the West German state of Hesse, where the environmental party which likes to be known as the Greens failed to repeat its successes of recent years and, unexpectedly, a coalition of the Christian Democrats and the Free Democrats is likely to form the next government. Does that mean that environmentalism is dead? Or is it merely that the environment now belongs to the middle classes?

There seems no question that it is the well-to-do who call the shots in the protection of the environment. They, after all, are those who have risen above the struggle for survival to a state in which they can regard smokestacks as an offence against nature and not a sign that there are jobs to be had, or regard tigers with affection and not fear. But how can it then be possible to make environmental policies for the world as a whole, given its diversity? This was the underlying but mostly unspoken theme of a Ditchley (north of Oxford) conference two weeks ago. In the protection of the environment, whose interests are at stake, those of the poor or of the rich?

There is a great deal in the belief that what would be generally considered as a proper concern for the natural environment is largely a function of prosperity. Only in the very richest countries, Japan and the United States (in some order), does one ever see people walking in the open with cotton gauze across their mouths in the belief that breathing unfiltered air will otherwise kill them. There are many other poorer parts of the world, Timbuctoo or London (England), where the same danger is more immediate. The general rule seems to be that a proper concern for the environment is being recognized for what it always has been, a kind of luxury.

That is not all that remarkable, as modest reflection will confirm. Not merely is there some kind of empirical correlation between gross national product *per capita* and national consciousness of environmental protection (in a sense, to be sure, that needs to be defined), but the same principle appears to apply at the level of the small group and even of the individual. Who but the well-to-do have the leisure, the surplus energy and the letter-writing capability to rise up in the defence of endangered species, landscapes or buildings?

Nor should the habitual defence of the natural environment predominantly by the wealthy, national or individual, be the occasion for ridicule and scorn. On the contrary, as the past few centuries of European history have vividly shown, the traditions of European art now widely revered were fostered and kept alive by the indulgence of the rich. Is Rembrandt the less to be respected because his customers were necessarily substantial people?

The flaw in this defence of the patronage of art (and even of mediaeval science) by those well-off enough to be able to match their luxuries to their tastes is that it is also a licence for fickle sponsorship. Wits say that the decision of the US Central Intelligence Agency to found *Encounter* magazine (which it did, although the two organizations have since parted company) shows how great things can stem from squalid motives; unhappily, the agency was probably up to other tricks at the same time. And the maxim that the rich know best, or that the market will decide, is generally distrusted where the natural environment is in question. The economists' explanation is that market preferences do not add up to a representation of the public good.

Some issues are nevertheless a considerable challenge to the notion that environmental protection is a self-evidently prudent policy for everybody. Take, for example, the species of whales still extant but much diminished in number. Like people, they sit at the top of elaborate food-chains which are largely separate from each other. The armchair ecologists' dictum that everything hangs together notwithstanding, it is hard to see what ecological harm would come to people if whales were promptly extinguished. This, apparently, is what the Icelandic, Japanese and Soviet members of the International Whaling Commission believe. Their more philosophical opponents on whaling rightly take the view that it would be a global tragedy if it were never again possible to go to sea in the hope of seeing one of these great creatures break the water. But that, of course, is a kind of luxury.

The more immediate issue is that of the tropical rain-forests, of which most of those who are able to purchase popular magazines have heard. There is a possibility that the steady destruction of these huge reservoirs of carbon will add to the greenhouse problem. Certainly, there is a

high chance that, once removed, the soil beneath will quickly erode. People are probably also on sure ground in saying that, in and beneath these dense canopies, there are probably uncounted species whose genetic constitution would be of great value at some point in the future. The case for not destroying the world's tropical forests, while not quite as powerful as that for not staging nuclear warfare, is very strong. So why the difficulty?

Most tropical forests have survived the past century because they are where the indigenous populations are so poor that they have until recently lacked the means to cut down the trees. And that is not a remarkable coincidence: rain-forests may be splendid ways of supporting a diversity of species but, left alone, they are not particularly hospitable to people. That may be why the rain-forest now most threatened, that of the Amazon, is within the territory of the largest but still, in *per capita* terms, the poorest of the developing countries. How can it hope to pay off its debts to the Western banks if it cannot sell a few cubic kilometres of irreplaceable hardwood for what the market will yield?

This, and in many other places, is where the contradictions of environmental protection arise most sharply. The most serious of those who would protect the environment, who make great sacrifices in their own lives to avoid damage to some special natural phenomenon, are prone to issue prescriptions that others should do likewise. Within national frontiers, the issue can be settled by the democratic process. Across frontiers, there is no means but rhetoric by which the rich can persuade the poor to do the decent thing.

Or not quite. The logic of international environmental protection and conservation is that, if it is the rich who want to keep things as they are, and the poor who need to change them for the sake of survival, the rich should compensate the poor for their restraint. In short, there is no basis for international policies on environmental protection that do not begin from the assumption that the policy-objectives are luxuries, are not to be scorned on that account and that their accomplishment requires a welding together of policies on environmental protection and foreign aid that has not so far been attempted. Unfortunately, to tell from the international rhetoric of the environment, the rich have calculated that exhortation will be cheaper.

John Maddox

Materials science

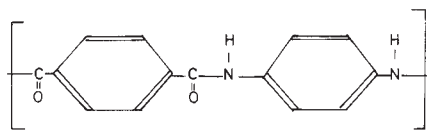
Synthetic polymer self-assembly

Paul Calvert

THE normal way to make a synthetic fibre is to polymerize material as an emulsion or suspension to form a powder, then to melt or dissolve that powder, squirt the liquid through a spinneret and draw it out into a strong fibre. In contrast, natural fibres such as collagen, cotton and wool seem to associate directly into fibres as they are polymerized. Certainly they are not melt-processed. So far there have really been no synthetic polymers that order in this way, but H.S. Yoon, on page 580 of this issue, now describes such self-association in an aromatic polyamide, a close relative of the commercial fibre Kevlar.

Mainstream polymer science is concerned with the behaviour of linear flexible macromolecules, mostly close relatives of the commercial polymers developed between 1930 and 1960, including polyethylene, polystyrene, aliphatic polyamides (nylons) and a dozen or so others. The molecules are very flexible, with weak forces between the chains, so that the polymer normally has the form of a random walk in three dimensions. To make a strong fibre it is necessary to pull the chains out straight and align them parallel to the fibre axis. This is achieved during drawing when the polymer is warmed and stretched to several times its original length.

During the 1970s the Dupont company developed aromatic polyamides such as poly-*p*-phenylene terephthalamide (see figure) which behave quite differently from the flexible polymers. The chains

Poly-*p*-phenylene terephthalamide.

have very little conformational freedom because both the phenyl group and the amide link are essentially planar and rigid. They can coil up, but only by a series of small changes in direction. It was found that if the solvent conditions are correct, these polymers form liquid-crystalline solutions where the molecules spontaneously align parallel to one another. These aligned solutions can be spun to strong fibres without a separate drawing stage. These and other liquid-crystalline polymers have been the subject of intense investigation over the past few years, but polymer science has not yet absorbed stiff chains and flexible chains into one seamless overall treatment.

Collagen is hydrophilic and bristles with hydrogen bonds that link chains together

into threefold helices, resulting in a very stiff structure. To form fibrous tendons it is necessary to export the polymer from the cell to the deposition site. The polymer is formed as procollagen, a short-chain soluble form with bulky terminal sections that stop the chains aggregating. Outside the cell the end sections are enzymatically removed and the chains form triple-helical rods of tropocollagen, 260 nm long. These rods assemble into rather weak fibres and then crosslink to form strong, high-molecular-weight, collagen fibres.

The significance of the new work of Yoon is then that the fibres form by a process very similar to the self-assembly of natural polymers. The polyphenylene terephthalamide system has much in common with silk or collagen — it is a stiff

chain with strong interactions between chains and strong solvent interactions. Yoon's most important result is that the polymer forms a metastable gel structure when polymerized in dimethylacetamide, and that this then slowly polymerizes further and aggregates into highly oriented fibres.

In discussing natural materials, it is tempting to ascribe all the control and self-organizing effects to powerful enzymatic mechanisms, because similar effects are not usually observed in synthetic materials. It may be that we are often also looking at the wrong types of materials. Flexible-chain polymers are relatively easy to characterize because they can be melted and dissolved. As polymer science spreads out into the less tractable polymers, we may find several phenomena which have previously been observed only in nature. Whether that will teach us more about polymers or more about nature remains to be seen. □

Paul Calvert is in the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

Superconducting ceramics

Temperatures rise higher still

M. Strongin, D.O. Welch and V.J. Emery

IN a recent News and Views report¹, we described the remarkable discovery of superconductivity near 40 K in the oxides (La, X)₂CuO₄ with X = Ba, Sr, Ca. The rapid pace of events made that report obsolete even before it was published: Chu and co-workers had reported² that the critical temperature, T_c , had more than doubled to 93 K in another yttrium-based oxide, Y-Ba-Cu-O. Even as that work appeared in *Physical Review Letters* on 2 March, other laboratories had confirmed the results and shortly afterwards it was shown that many other elements could be substituted for yttrium to give values of T_c in the neighbourhood of 90 K. Indeed, one of these, La-Ba-Cu-O, contains the same elements as the original material discovered by Bednorz and Müller³, which set this sequence of events into motion. Although this higher T_c phase contains the same elements, it has a different composition and crystal structure. A special session devoted to these developments was held on 18 March in New York at the annual meeting of the American Physical Society (see pages 432–433 of last week's issue).

The degree to which these discoveries have excited the condensed-matter physics and materials-science communities was borne out by the extraordinary character of the session, which was attended by a vocal, enthusiastic audience of more than 2,000 scientists, many of

whom viewed the proceedings by closed-circuit television. In an appreciation of the historical significance and wide interest in the proceedings, a videotaped record has been prepared and copies can be obtained from the American Physical Society.

The session consisted of about 60 presentations and lasted from 7.30 p.m. until the early hours of the morning, with some diehards holding out until 6.00 a.m. After introductory remarks, the session began with presentations and discussions by a panel of experimentalist 'pioneers' in the field consisting of K.A. Müller (IBM Zurich), S. Tanaka (University of Tokyo), C.W. Chu (University of Houston), Z.X. Zhao (Chinese Academy of Sciences) and B. Batlogg (AT & T Bell Laboratories). By this time all the groups had prepared the yttrium-containing oxides, and the superconducting phase (the original report of 93 K was for a multiphased sample) had been identified as YBa₂Cu₃O_m with $m \sim 7$. Impressive strides had also been made in fabricating the superconductors using ceramic processing techniques, and Batlogg evoked an enthusiastic response from the large crowd by displaying a roll of pliable tape made from unfired material. The initial presentations of experimental findings were followed by short talks by P.W. Anderson (Princeton University), V. Kresin (Lawrence Berkeley Laboratory), A.J. Freeman (Northwestern University),

D. Papaconstantopoulos (Naval Research Laboratory) and M. Schluter (AT&T Bell Laboratories), who discussed various theoretical ideas, ranging from state-of-the-art electronic-band structure calculations and estimates of T_c based on the usual electron-phonon interactions to more exotic pairing mechanisms based on charge fluctuations, acoustic plasmon exchange and resonating valence bonds. Although it had been shown that the early 'low- T_c ' materials ($T_c = 40$ K) can be accommodated in the usual picture of superconductivity produced by strongly coupled electron-phonon interactions, there is a growing feeling that the 'high T_c ' demands something more.

Following the initial panel presentations, various research groups presented 5-minute summaries of a wide range of theoretical and experimental results. Rapid progress has been made in determining the crystal structure of the so-called '1-2-3' phase, with the focus on studies of $\text{YBa}_2\text{Cu}_3\text{O}_m$ and, although there is still some debate on the exact number and location of oxygen atoms in the unit cell, several features of the structure have emerged fairly clearly. As in the case of the $(\text{La}, \text{X})_2\text{CuO}_4$, the new phase is also a variant of the perovskite structure, with layers of copper and oxygen playing a prominent role. However, whereas $(\text{La}, \text{X})_2\text{CuO}_4$ has perovskite-like layers of LaCuO_3 separated by two rock salt-like LaO layers, the new structure has a tripled simple perovskite unit cell with three planes containing copper and oxygen. In one version (from AT&T Bell Laboratories), the electronic bands are two-dimensional in two of the planes, but one-dimensional in the third, and the suggestion has been made that low-dimensionality is an important factor in the high values of T_c . Despite the remaining uncertainties concerning the exact oxygen content and location, it seems clear that the main features important for superconductivity must be contained in the properties of the CuO_2 layers, as several groups find that many other rare-earth elements, including, surprisingly, the strongly magnetic gadolinium, can substitute for yttrium with comparable values of T_c .

As we write this article, one of the chief uncertainties about the new oxide superconductors is the possible range of values of the critical current density. Values have been reported (J. Gavaler, Westinghouse Research Laboratories) of the order of 10^4 A cm^{-2} in magnetic fields of a few tesla, but this is still much smaller than required for high-field magnets. Whether this is an intrinsic problem of this class of superconductors is still unknown, but it is very likely that better processing will lead to higher critical currents. In any event, it is much more certain that the critical magnetic fields are immense, greater than a million gauss; although this is somewhat

marred by reports of substantial anisotropy when the field is parallel or perpendicular to the planar structure. At present the study of the new high- T_c superconductors is less than a month old. Already, because of better processing, critical temperatures up to 120 K are being reported (H.L. Luo, University of California, San Diego and C. Politis, KFK, Karlsruhe), and there are reports, including those from the groups at Berkeley and Houston, of irreproducible sharp resistance drops at temperatures up to 240 K.

Even as this article goes to press the *New York Times* on 28 March reported that workers at Wayne State University have observed the 'a.c. Josephson effect'

in samples of Y-Ba-Cu-O at 240 K. This provides further evidence that there are small amounts of some new superconducting phase at this temperature. It is clear that many surprises are still in store for the adventurous, and few will contest the value of the great contributions made by Bednorz and Müller and by Chu and his co-workers. It is a great tribute to creativity in experimental science. □

1. Strongin, M., Welch, D.O. & Davenport, J.W. *Nature* **325**, 664-665 (1987).
2. Wu, M.K. *et al.* *Phys. Rev. Lett.* **58**, 908-910 (1987).
3. Bednorz, J.G. & Müller, K.A. *Z. Phys.* **B64**, 189-193 (1986).

M. Strongin, D.O. Welch and V.J. Emery are at the Brookhaven National Laboratory, Upton, New York 11973, USA.

Gamma-ray astronomy

Surprises in the skies

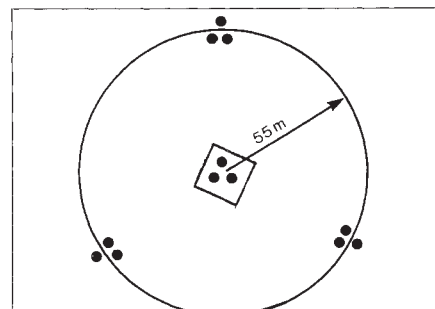
A.A. Watson

THE list of astronomical objects known to emit TeV (10^{12} eV) γ -rays is growing rapidly and has been further extended by the claim by A.R. North and colleagues on page 567 of this issue¹ that Vela X-1 is such a source. This work is particularly interesting because the X-ray pulsar period, 283 s, of Vela X-1 is more than 70 times longer than that of 4U0115+63, the previous slowest X-ray binary pulsar detected at TeV energies. The reported TeV luminosity of Vela X-1, at 2×10^{34} erg s^{-1} , is about 10 times lower than for other known sources. As it lies only 1.4 kpc from the Earth, the implication is that many similar objects will be detected as telescope sensitivities improve. Furthermore, it is the first TeV source in which the matter accreted around the neutron star arises from a stellar wind blown off the massive OB-companion star, rather than from Roche lobe overflow.

The catalogue of TeV sources includes systems as diverse as the powerful radio galaxy Cen A, the isolated pulsars in the Crab and Vela supernovae remnants, and several binary X-ray pulsar systems including Cygnus X-3 and Hercules X-1. The number of TeV γ -ray photons emitted by such sources is typically less than 10^{-6} m^{-2} s^{-1} , so that even a 1- m^2 satellite-borne detector can record only a handful of such photons per year. But when they enter the atmosphere these γ -rays generate a small electromagnetic cascade, the charged particles of which radiate Cerenkov light as they pass through the air. The light pool generated at ground level by the cascade covers an area of about 10^5 m^2 over which the photon density is reasonably constant. Several Cerenkov photons per square metre are produced in the visible region of the spectrum and are readily detected against the night sky background on clear moonless nights

using photomultipliers mounted at the foci of large mirrors. The earliest attempts to detect very high energy γ -rays from celestial sources used single-searchlight mirrors, but purpose-built mirror assemblies have begun to feature in modern telescopes. The telescope used in the new study¹ of Vela X-1 consists of three separated clusters, each with three 1.5-m diameter mirrors (see figure). The mirrors were used to track the source and the time of threefold coincidences between signals in the mirrors of each cluster were recorded with a resolution of 0.1 μs .

The evidence that Vela X-1 is an emitter of TeV γ -rays is twofold. First, the source was observed for a total of 26.6 h on 11 nights between 2 April and 10 May 1986. Power-spectrum analysis of the coincidence times shows a peak at a period of 282.805 s that lies near the middle of the range of X-ray pulsar periods² found for this source since its first detection. The time-averaged pulsed flux during the observations is $(2.0 \pm 0.4) \times 10^{11}$ photons cm^{-2} s^{-1} and the probability of this arising by chance is estimated as 6.6×10^{-4} . The observations of North *et al.* were made between phases 0.23 and 0.91 of the



Schematic representation of the Potchefstroom TeV γ -ray telescope, showing four clusters of mirrors. Only three of the clusters were used in the Vela X-1 observations.

8.9-day orbital period and no evidence was found of orbital modulation. Second, in addition to the steady output they detected a very strong burst of emission, lasting about 19 min, on 4 May 1986. For about 1.5 min during the outburst the mean count rate increased by 100 per cent above the background rate of about 0.8 Hz. The enhancement occurred in each of the three independent telescopes and corresponds to a d.c. excess of $8 \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$. North *et al.* also observed periodic emission at 282.805 s during the four pulsar rotations following the outburst. The combined probability for the pulsed outburst and the d.c. emission resulting from chance is 1.1×10^{-5} . Taken together, then, these data make a convincing claim for Vela X-1 to be added to the catalogue of TeV sources.

An additional, remarkable feature of the observations of North *et al.* is that the outburst occurred just over 4 h after the X-ray source had entered eclipse by its companion star. A similar observation at about 0.6 TeV has been reported³ from Hercules X-1 in which pulsed emission was detected for more than 1 h after the beginning of the 6-h eclipse. The γ -ray emission from Hercules X-1 can be accounted for by postulating² neutral pion production in the surrounding matter by a proton beam accelerated near the neutron star. To explain TeV γ -ray emission while the beam source is in eclipse also requires the proton beam to be steered⁵ in the weak field of the companion star. North *et al.*¹ invoke the same mechanism in the case of Vela X-1 and, if detailed analysis substantiates their conjecture, it seems probable that there is now direct evidence for high-energy proton beams both in Vela X-1 and in Hercules X-1. The protons that escape from such systems will form part of the galactic cosmic-ray population; thus, in addition to the importance of the new findings for the physics of compact objects, X-ray binaries may be involved in solving the problem of the origin of cosmic rays.

As there have been claims^{6,7} that Vela X-1 also emits γ -rays above 1 PeV (10^{15} eV) the new work¹ raises to three (Cygnus X-3 and Hercules X-1 being the others) the number of X-ray binary systems that emit across the range 1 TeV–1 PeV. The pulsars in those systems have periods which differ by more than four orders of magnitude, from the 12.5 ms of Cygnus X-3 (ref. 8) to the 282 s of Vela X-1. It is far from clear how an accretion-fed pulsar can accelerate protons to create the γ -ray-producing cascades. Further experimental measurements are needed; these may come rapidly as, in addition to the authors of ref. 1, a group from Durham University is now patrolling the southern skies with a new telescope recently established at Narribri, New South Wales. For PeV work, efforts of the Adelaide⁶ and

Chacaltaya⁷ groups will soon be supplemented by a joint Bartol Research Foundation/Leeds University 100-TeV telescope at the US South Pole base. This device will allow sources such as Vela X-1 and Cen X-1, another wind-driven X-ray binary, to be observed simultaneously at two very different energies: no doubt these amazing objects hold further surprises. □

1. North, A.R., Raubenheimer, B.C., de Jager, O.C., van Tonder, A.J. & van Urk, G. *Nature* **326**, 567–569 (1987).

2. Joss, P.C. & Rappaport, S.A. *A. Rev. Astr. Astrophys.* **22**, 537–592 (1984).
3. Gorham, P.W. *et al. Astrophys. J.* **308**, L11–L15 (1986).
4. Eichler, D. & Vestrand, W.T. *Nature* **318**, 345–347 (1985).
5. Gorham, P.W. & Learned, J.G. *Nature* **323**, 422–424 (1986).
6. Protheroe, R.J., Clay, R.W. & Gerhardy, P.R. *Astrophys. J.* **280**, L47–L50 (1984).
7. Suga, K. *et al. Proc. Workshop on Techniques in UHE Gamma Ray Astronomy, La Jolla* (eds Protheroe, R.J. & Stephens, S.A.) 48–53 (University of Adelaide, 1985).
8. Chadwick, P.M. *et al. Nature* **318**, 642–644 (1985).

A.A. Watson is Professor of Physics at the University of Leeds, Leeds LS2 9JT, UK.

Cell biology

Cyclins in meiosis and mitosis

Andrew W. Murray

IN A recent paper¹, Swenson, Farrell and Ruderman demonstrate that a protein from the surf clam egg, whose abundance fluctuates during the embryonic cell cycle, can induce meiosis when it is synthesized in frog oocytes. This finding represents a flowering of molecular physiology: the combined application of molecular biology, manipulation of cells and crude *in vitro* systems to shed light on problems in cell biology *in vivo*. In this case the approaches used are the molecular analysis of the pattern of protein synthesis in embryonic cells and studies of the intracellular factors that cause cells to enter mitosis and meiosis. The contribution of the humble surf clam to this synthesis is a reminder that unfashionable, genetically intractable organisms can still end up on the cover of fashionable journals.

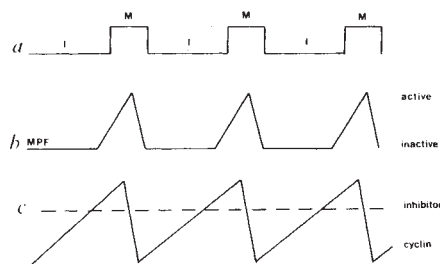
Two lines of evidence from cell physiology experiments led to the idea that there are inducers of mitosis and meiosis. The first is that frog oocytes induced to enter meiosis by progesterone produce an activity called maturation promoting factor (MPF, see ref. 2 for a review), which stimulates the maturation of oocytes that had not been exposed to progesterone. The second comes from cell fusion experiments in which mitotic HeLa cells were fused to cells in other parts of the cell cycle. In all cases the interphase nucleus is induced to enter mitosis. These two approaches were combined when it was shown that somatic cells in mitosis also contain MPF, suggesting that the same factors induce both mitosis and meiosis. Despite intense investigation and the development of *in vitro* systems in which the addition of crude MPF stimulates mitotic events such as nuclear envelope breakdown, chromosome condensation and spindle formation^{3,4}, progress in the molecular characterization of MPF has been limited. Even in *in vitro* systems, it is not known how many intermediate steps exist between the addition of MPF and the observed changes.

The molecular approach has taken place mainly at the Marine Biological Lab-

oratory in Woods Hole, Massachusetts. It stemmed from the knowledge that protein synthesis is required for the appearance of MPF and for the occurrence of mitosis in embryonic cells. Because embryonic cells divide without increasing in size and have considerable stockpiles of most of the components that are duplicated during the cell cycle, it seemed likely that the proteins required for the induction of mitosis might constitute an appreciable fraction of embryonic protein synthesis. Hunt and colleagues examined⁵ the pattern of protein synthesis in sea urchin eggs in the first few cell cycles after fertilization. They found that one of the proteins, which they named cyclin, increases in abundance during each interphase and is destroyed at each mitosis. Clams contain two cyclins, at least one of which has a very similar amino-acid sequence to sea urchin cyclin (ref. 1 and T. Hunt, unpublished results). In both sea urchins and clams⁶ the translation of cyclin messenger RNA is activated at fertilization, which corresponds to the resumption of the cell cycle after a long period of interphase arrest. In the clam, both meiotic divisions occur after fertilization and cyclin levels rise and fall during both divisions. Arresting embryos in either mitosis or meiosis stabilizes cyclins for long periods. It was tempting to speculate that cyclin accumulation causes cells to enter mitosis or meiosis and that cyclin destruction is necessary for return to interphase.

Swenson *et al.* have now been able to link molecular biology and cellular physiology by showing¹ that the transcript of the cloned gene encoding clam cyclin A induces the first meiotic division when injected into oocytes of the South African clawed toad (*Xenopus*). The authors cannot tell whether the second meiotic division also occurs. As they admit, their inability to demonstrate that cyclin induces meiosis II or mitosis weakens the idea that cyclin is a general inducer of mitosis and meiosis. This problem is exacerbated by the fact that meiosis in *Xenopus* is induced by the hormone progesterone, which has

no role in the mitotic cycle. At worst, cyclin could be a component in the response to progesterone, such as the regulatory subunit of cyclic AMP-dependent kinase (which can induce meiosis when injected into *Xenopus* oocytes). In addition, the mechanism of MPF activation in meiosis I seems to be different from that in meiosis II or in mitosis. In both frogs and clams meiosis I is the only division that can be induced in the absence of protein synthesis. In the clam meiosis I can occur in the absence of either newly synthesized or pre-existing cyclin A (ref. 1). The simplest interpretation of these results is that MPF in meiosis I can be activated in the absence of cyclin, whereas the activation of MPF in meiosis II and mitosis requires cyclin synthesis. If this view is correct it implies that cyclin is an activator of MPF, rather than MPF itself, and that the ability of cyclin to induce meiosis I is not necessarily a good test for its role in meiosis II or mitosis. Despite this caveat, the original



Schematic representation of the embryonic cell cycle. *a*, Regular alternation of interphase (I) and mitosis (M); *b*, appearance and disappearance of MPF activity, which causes cells to enter and leave mitosis, respectively; *c*, levels of cyclin and an inhibitor that inactivates MPF. MPF activity appears when cyclin levels exceed that of the inhibitor and disappears when cyclin levels fall below that of the inhibitor.

demonstration that fluctuations in cyclin abundance parallel progress through the mitotic cell cycle argues that cyclin is important in the control of mitosis.

Like cyclin, MPF disappears as cells exit from mitosis or meiosis. Does the decline of MPF correspond to the transition between metaphase and anaphase and, if so, is a preceding destruction of cyclin necessary to allow conversion of MPF to an inactive form? The figure shows a scheme where MPF is activated by cyclin and inactivated by an inhibitor that is present throughout the cell cycle. Such a scheme fits with those derived by Nurse and colleagues from genetic studies on *Schizosaccharomyces pombe*. They have shown that the *cdc2* gene product, which is required to enter mitosis, is activated by the *cdc25* gene product and inhibited by the *wee1* gene product (see ref. 7 for a review). This analogy makes *cdc2* and *cdc25* appealing candidates for the genes encoding yeast MPF and cyclin, respectively.

Why should the activity of cyclins be controlled by a means as drastic as pro-

teolysis? One answer is that the irreversibility of their inactivation ensures that the cell cycle runs in only one direction. The second class of explanation is that cyclin is a component of a cell-cycle clock (see figure). At the beginning of the cell cycle the activity of the putative inactivator of MPF is high, cyclin is absent and MPF remains in its inactive form. During interphase, cyclin accumulates until eventually the rate of MPF activation exceeds that of inactivation. As a result high levels of MPF accumulate and the cells enter mitosis. If these levels of MPF lead to the destruction of cyclin, MPF will be destroyed by the inactivator and the cells will return to interphase. The existence of cycles of this type would help to dispel the long-standing arguments between those who see the cell cycle as a series of events that run in a fixed sequence and those who view it as the reflection of an underlying oscillator. The gradual accumulation of cyclin could be both the step required for MPF activation (and, ultimately, the destruction of cyclin and inactivation of MPF) and the periodic event that determines the cycle length of an oscillation in MPF activity. In *Xenopus*, the oscillation of MPF is unaffected by treatments that perturb the nuclear cell cycle⁸ and continues even in the total absence of nuclei⁹.

Although embryonic cell cycles are short and synchronous, somatic cell cycles are long and variable in length. This difference can be explained if the cell-cycle clock in embryonic cells starts at the end of mitosis and that of somatic cells starts at a variable time after the end of the preceding mitosis. In somatic cells the ability to start a clock that counts down to mitosis could be made dependent on attaining a sufficient cell size, obtaining adequate nutrient supplies or the completion of DNA synthesis. All these controls have been demonstrated⁷ in *S. pombe*. In addition, biochemical parameters, such as carbon dioxide production and enzyme activity, continue to fluctuate in a periodic manner even when the nuclear cycle of *S. pombe* is arrested¹⁰, suggesting that the existence of cytoplasmic clocks is not restricted to embryonic cell cycles. □

1. Swenson, K.I., Farrell, K.M. & Ruderman, J.R. *Cell* **47**, 861-870 (1986).
2. Kirschner, M., Newport, J. & Gerhart, J. *Trends Genet.* **1**, 41-47 (1985).
3. Lohka, M.J. & Maller, J.L. *J. Cell Biol.* **101**, 518-523 (1985).
4. Miale-Lye, R. & Kirschner, M. *Cell* **41**, 165-175 (1985).
5. Evans, T. *et al.* *Cell* **33**, 389-396 (1983).
6. Rosenthal, E.T., Hunt, T. & Ruderman, J.R. *Cell* **20**, 487-494 (1980).
7. Hayles, J. & Nurse, P. *J. Cell Sci. Suppl.* **4**, 155-170 (1986).
8. Gerhart, J., Wu, M. & Kirschner, M. *J. Cell Biol.* **98**, 1247-1255 (1984).
9. Wasserman, W.J. & Smith, L.D. *J. Cell Biol.* **78**, R15-R22 (1978).
10. Creanor, J. & Mitchison, J.M. *J. Cell Sci.* **86**, 207-215 (1986).

Andrew W. Murray is in the Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA.

Which star went bang?

AN amusing but significant sub-plot in the unfolding saga of supernova 1987A has arisen from the inability of astronomers to decide which star was its progenitor. The initial favourite, Sanduleak (Sk) -69 202, was reportedly detected again a few days after the explosion. Now astronomers are suggesting, albeit with qualification, that this was after all the star that erupted.

Before the supernova occurred, Sk -69 202 had been observed to have the brightness and spectral characteristics of a hot blue supergiant. The positional coincidence of this star and the supernova, combined with the theoretical belief that such a star can end its nuclear-burning lifetime in just such a conflagration, made it an ideal progenitor candidate. It was disconcerting, therefore, when C. Gry and other European astronomers working with the International Ultraviolet Explorer (IUE) satellite detected Sk -69 202 six days after the supernova was first seen.

The ultraviolet behaviour of the supernova itself is somewhat puzzling. At optical wavelengths it has many characteristics of a type II event. In contrast, the ultraviolet signatures are more similar to type I supernovae, thought to originate in more evolved stars of smaller mass. There is in fact good reason to suppose that SN 1987A is a peculiar type II event, not least from novel theoretical considerations that will be discussed here next week. But in either interpretation it was unexpected that the far ultraviolet flux recorded by IUE decreased as exceedingly rapidly as it did. The happy consequence was that the underlying stellar spectra became discernible, if only at these restricted wavelengths.

The combination of careful studies of previous photographic plates of this portion of the Large Magellanic Cloud with the most recent IUE spectra has now led to peculiar difficulties. The plates show that there were three stars aligned in a north-westerly direction over an angular distance of five arc seconds, Sk -69 202 being the inner star. The IUE spectra now reveal two stars. One, a B0 supergiant, can be matched with the northernmost star on the optical plates. But the ultraviolet spectrum of the southern star, according to N. Panagia and colleagues in a recent preprint, conforms more closely to the southernmost optical star than to Sk -69 202, those stars being only one second of arc apart on the plates. Such angular distances are at the limits of the capabilities of IUE and, what is more, uncertain corrections for interstellar reddening form a crucial part of the argument. Nevertheless, according to Panagia *et al.*, "it looks as if Sk -69 202.... has become supernova". One intriguing question now is whether IUE astronomers in the United States will concur with such a revision.

Philip Campbell

Optical technology

Novel optical properties of liquid-crystal polymers

George Attard and Graham Williams

ORIGINALLY a curiosity, liquid crystals were synthesized and studied primarily for academic reasons¹ until, in the late 1960s, it became apparent that they were ideal for use as active media in display devices. This development, together with the increasing variety of applications involving liquid-crystalline materials, is one of the great success stories of modern science and technology. More recently, macromolecular liquid-crystal polymers have shown a combination of electroactive and photoactive properties with the special physical properties of liquid crystals and synthetic polymers. I. Cabrera and V. Krongauz report on page 582 of this issue² one liquid-crystal polymer which they show exhibits thermochromism — the property of changing colour progressively with temperature. This is one of several optical properties of liquid crystals that are of great interest to those working on the technology based on 'photonics' for the twenty-first century.

The liquid-crystalline phase (mesophase) is one that occurs between liquid and solid, in which molecules flow but also have orientational order. Low-molar-mass

liquid crystals are commonly based on a monomer consisting of an anisometric (cigar-, lath-, or disk-like) head group and a flexible 'tail'. Anisotropy arises as a result of the tendency for the head groups to align with their long axes (cigar- and lath-like), or planes (disk-like) parallel. This long-range anisotropy, encompassing thousands of molecules, is responsible for the birefringence observed in liquid-crystalline phases.

In the late 1970s, a new class of liquid crystals, liquid-crystal polymers, were synthesized and investigated. These are now of great interest because of their recently demonstrated optical and electrical properties³⁻⁷. In these, the anisotropic head groups (moieties) are connected to a polymer chain, either as part of that chain (main chain), hanging off the chain (pendant side-chain) or a combination of the two. The figure illustrates typical arrangements. Copolymers (synthesized from more than one monomer) may be made in which some side chains (mesogenic) confer the tendency to form liquid-crystalline phases and other groups confer some desirable property. Indeed it

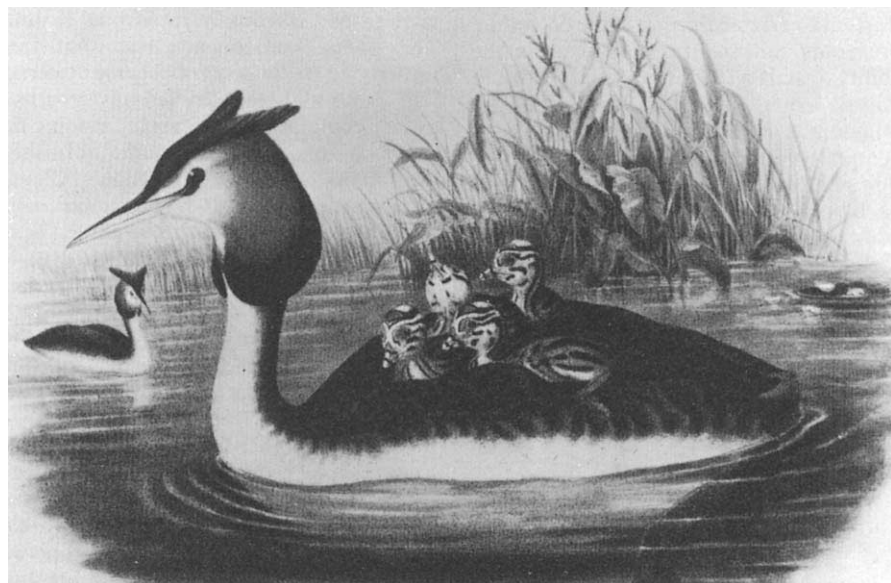


Schematic representation of the rich molecular architecture of liquid-crystalline polymers, showing side-chain and main-chain types with disk- and cigar-type moieties.

is common for copolymers to assume a multifunctional character. Thus, groups that are thermochromic, pleochroic (in which the colour intensity depends on orientation) or photochromic (chemical changes brought about by light alter the liquid-crystal colouring), and other functional groups, may be included in the copolymer. A major advantage of multifunctional copolymers is that the problem of phase separation that occurs in mixtures of mesogenic monomers and functional monomers is avoided.

The principles that govern liquid-crystal formation in low-molar-mass mesogens can be extended to the polymeric systems. The latter, being hybrid materials, also possess polymer characteristics: namely, good film-forming properties, viscoelasticity (elasticity in flow) and the formation of glassy phases at lower temperatures. Side-chain liquid-crystal polymers are of special interest as they exhibit the electro-optical behaviour⁸ found in low-molar-mass mesogens. Consequently, many of the applications envisaged for these materials overlap with those for which low-molar-mass liquid crystals are already used. Because of their high viscosity, however, polymers are not suitable for use in active display devices (except possibly as solutes to modify the properties of low-molar-mass mesogens). The high viscosity and glassy nature of the mesophases formed by these materials is well exploited when they are used as media for laser-addressed reversible information storage^{3,4,9}.

Many other applications are envisaged for multifunctional liquid-crystalline polymers, including their use as specialized optical components (Fresnel gratings, for example), and as waveguides in integrated optical circuits. Furthermore, their non-linear optical properties are of particular interest in view of the current research into non-linear materials for use in optical computing systems. The efficiency of frequency doubling of a laser beam is substantially increased in organic



This picture of the great-crested grebe is on display at an exhibition of the works of the nineteenth-century zoologist, ornithologist and publisher John Gould at the Natural History Museum, Cromwell Road, London from 8 April to 27 September 1987. Gould produced more than 40 imperial folio volumes containing about 3,000 hand-coloured lithographs showing birds of every continent except Africa. The exhibition brings together a collection of Gould's work from all over the world, and offers the opportunity to view watercolours, paintings, letters and scrapbooks that he compiled when working on the *Birds of Great Britain*. The exhibition is organized by the Royal Society for Nature Conservation together with the Osberton Trust to support the British Wildlife Appeal and to celebrate the 75th anniversary of the foundation of the society. □

systems if they form mesophases¹⁰. Polymers below their glass transition can retain an induced non-centrosymmetrical arrangement, which is essential for frequency doubling. This gives liquid-crystal polymers a major advantage over their monomer counterparts. Furthermore, polymeric systems exhibit high laser-damage thresholds that allow them to be used with powerful lasers, which would rapidly damage inorganic or organo-metallic materials.

An understanding of the structure, molecular dynamics and physical properties of mesogenic side-chain polymers, and of their behaviour in electric and magnetic fields, is slowly emerging as a result of X-ray, optical nuclear magnetic resonance and dielectric studies^{3,4,11}. But many questions remain unanswered and several aspects need clarification for the full technological potential of these new materials to be realized. Much needs to be done on controlling molar mass (high-molar-mass materials are only now being synthesized); the effect of the geometry of the main chain (stereoregularity) on the

material properties and the wide variety of chemical structures that could be incorporated into the polymers have yet to be explored. In addition, polymeric systems that form mesophases under the influence of solvents have received little attention¹¹ but are expected to give rise to some important developments in the formation of functional Langmuir-Blodgett films. □

1. Gray, G.W. *The Molecular Structure and Properties of Liquid Crystals* (Academic, New York, 1962).
2. Cabrera, I. & Krongauz, V. *Nature* **326**, 582-585 (1987).
3. Shibaev, V. P. & Plate, N.A. *Adv. Polym. Sci.* **60/61**, 173-252 (1984).
4. Finkelmann, N. & Rehage, G. *Adv. Polym. Sci.* **60/61**, 99-172 (1984).
5. Kalisky, Y. & Williams, D. J. *Macromolecules* **17**, 292-296 (1984).
6. Dobb, M. G. & McIntyre, J. E. *Adv. Polym. Sci.* **60/61**, 61-98 (1984).
7. Engel, M. *et al.* *Pure appl. Chem.* **57**, 1009-1014 (1985).
8. Araki, K., Attard, G. S. & Williams, G. *J. molec. Electron.* (in the press).
9. Coles, H.J. & Simon, R. *Polymer* **26**, 1801-1806 (1985).
10. Williams, D.J. (ed.) *ACS Symp. Ser. No. 233* (1983).
11. Attard, G. S. & Williams, G. *Chem. Br.* **22**, 919-924 (1986).

George Attard and Graham Williams are at the Edward Davies Chemical Laboratories, University College of Wales, Aberystwyth SY23 1NE, UK.

Forest ecology

Tree boundaries on the move

Peter D. Moore

JUST as Macbeth made an unfortunate error of judgment in assuming that Great Birnam Wood could never move to Dunsinane, so can the inherent mobility of forests easily be underestimated. The climatic fluctuations of the past two million years have resulted in dramatic migrations of forests, and recent research suggests¹ that some trees have achieved surprisingly rapid rates of movement using particularly mobile vectors, such as the now-extinct passenger pigeon. Not only are the means of travel somewhat obscure, so are the migration routes: new biochemical and cytological surveys show² that the origins of many tree populations, even the Caledonian forests of Scots pine in Scotland, are not as clear-cut as might be thought.

Migration rates of forests can be studied by the measurement of age distribution along transects at right angles to the forest edge³, but the forest front is not always easily discernible. A more common method that provides a broader view of movement is the collation of data from pollen diagrams. Movements of lines of equal pollen input (isopolls) through time can be assumed to reflect the movements of tree species⁴. Studies of the polewards movements of trees since the end of the last glaciation about 10,000 years ago demonstrate that migration rates were surprisingly rapid, often in excess of 300 m per year⁵. In North America, Webb⁶ calculates that the forest was advancing by

between 1 and 8 km per generation, depending on the tree species involved. Even more remarkable, as Webb points out, is the similarity between the rates of movement of tree species despite very different modes of dispersal. The migration rate of the white pine (*Pinus strobus*), for example, with its seeds dispersed by winds, is equalled by that of the heavy-fruited oaks (about 350 m per year)⁵.

This rapid rate of movement can only be achieved by heavy seeds if they are transported by a mobile vector. Although some mammals move and store seeds, their ability to transport viable propagules over long distances is limited. Many birds, such as jays, nutcrackers and titmice, hoard food but their capacity to relocate caches of nuts is prodigious⁶ and they are unlikely to carry them far beyond the forest margin. Migratory species are more likely to provide a suitable vector. Studies⁷ on the island of Surtsey, off Iceland, reveal that small passerines such as the snow bunting can carry viable seeds in their gizzards from mainland Britain about 1,000 km away. In North America, Webb suggests the passenger pigeon (*Ectopistes migratorius*) was a vector, but this bird did not carry and store seeds, so its effectiveness depended on the survival of seeds passing through the gut. She argues that because the crop of a passenger pigeon could hold up to 30 acorns the possibility of bird death, regurgitation or viable

passage was considerable, especially taking into account the vast flocks of these birds which once foraged and migrated over wide areas of North America. There is archaeological evidence for the presence of the passenger pigeon since early in the present interglacial period.

There were no passenger pigeons in the British Isles, yet the rates of advance of oaks and hazel were similar to those of North America at that time. Perhaps the wood pigeon (*Columba palumbus*) was the major vector on the British side of the Atlantic. Eastern European populations of this bird are at present clearly migratory⁸, although oceanic birds tend to be more sedentary. It is impossible to say, of course, how the species behaved 10,000 years ago.

When dealing with islands, there are further problems to be considered resulting from the lowered sea level during glacial times and the opportunities for warmth-demanding trees to find refugia in low coastal sites. If such currently submerged refugia existed, then the isopoll lines should provide evidence for them. There is indeed a suggestion of spread from isolated western locations in the present-day interglacial period for hazel and for alder.

An alternative approach in the search for roots has been supplied by the monoterpene and isoenzyme studies on the Scots pine (*Pinus sylvestris*) of Kinloch, Westfall and Forrest². The Caledonian race of this species is isolated from other populations, as the pine became extinct over much of England and Wales by about 3,000 years ago, although it survived longer in Ireland. Scots pine was an early postglacial invader of southern England, almost certainly as a result of migration across the land link with continental Europe, but it expanded in eastern Scotland only about 7,000 years ago. Was this invasion an extension of the English pine northwards, or did the Scots pine have a separate origin? The biochemical evidence suggests the latter, for the Scottish trees display little genetic affinity with those of mainland Europe. The distinctiveness of the Caledonian race suggests that it has been isolated for a long time and that its origin was in western refugia, perhaps in the region of the Hebrides. □

1. Webb, S.L. *Quat. Res.* **26**, 367-375 (1986).
2. Kinloch, B.B., Westfall R.D. & Forrest G.I. *New Phytol.* **104**, 703-729 (1986).
3. Leak, W.B. & Graber, R.E. *Ecology* **55**, 1425-1427 (1974).
4. Huntley, B. & Birks, H.J.B. *An Atlas of Past and Present Pollen Maps of Europe: 0-13,000 Years Ago* (Cambridge University Press, 1983).
5. Davis, M.B. in *Forest Succession: Concepts and Applications* (eds West, D.C., Shugart, H.H. & Botkin, D.B.) 132-153 (Springer, New York, 1981).
6. Shettleworth, S.J. *Scient. Am.* **248** (2), 86-94 (1983).
7. Fridriksson, S. *Surtsey: Evolution of Life on a Volcanic Island* (Butterworths, London, 1975).
8. Cramp, S. *et al.* *The Birds of the Western Palearctic Vol. 4* (Oxford University Press, 1985).

Peter D. Moore is in the Department of Biology, King's College London (KQC), Campden Hill Road, London W8 7AH, UK.

Neurosensory signalling

Biochemical puzzles about the cyclic-GMP-dependent channel

Meredith L. Applebury

PHOTORECEPTOR cell signalling results from a transient hyperpolarization of the plasma membrane — a process that is achieved by the closing and reopening of a cyclic (c) GMP-gated channel. An extremely sensitive system for detecting light and hydrolysing cGMP, the rhodopsin-mediated cGMP cascade depletes the cGMP-gating messenger and guanylyl cyclase replenishes cGMP (see figure). Dissecting this system in molecular detail to determine the sequence and nature of events is a continuing challenge¹. The key components of the receptor-mediated cGMP cascade, rhodopsin, G protein and cGMP-phosphodiesterase, are well characterized^{2,3}. Now, the channel itself has been identified biochemically in two laboratories. Alas, the results are baffling. Cook *et al.*⁴ have isolated and reconstituted a cGMP-gated channel composed of subunit(s) of relative molecular mass 65,000 (65K). And on page 600 of this issue⁵, Matesic and Liebman examine a channel with similar characteristics consisting of 39K subunit(s). How can these two apparently conflicting sets of data be interpreted?

The idea that an ion channel could be directly regulated by a cyclic nucleotide ligand was slow to gain acceptance. As early as 1978, Nicol and Miller reported⁶ that cGMP injected into a rod cell of the retina affected membrane conductances. Caretta *et al.* then demonstrated⁷ the presence of a cGMP-gated release of cations in rod outer-segment membrane preparations. But reservations persisted¹ until Fesenko *et al.* used⁸ patch-clamp methods to establish that a cGMP-gated conductance resides in the plasma membrane. Over the past two years, this channel has been the focus of intense research.

Physiologically, the cGMP-dependent conductance is a strange channel (see ref. 9 for a review). The cGMP binds allosterically to open the channel, and the concentrations measured at which it is 50 per cent effective ($K_{1/2}$) are 5–45 μM . The Hill coefficients (n) are 1.5–3.1, suggesting that there are 2–4 cooperative sites. The channel is permeable to both monovalent and divalent cations. Under physiological conditions, the light-sensitive current

shows almost no voltage dependence over the range in which the cell operates (–30 to –65 mV) and the unitary conductance is a record low, about 3 fS. These properties may be caused by transient blocks from divalent cations. With 100 mM Na^+ and Mg^{2+} and Ca^{2+} concentrations below 10^{-7} M, the channel has two conducting states, 25 and 8 pS, and an open time of

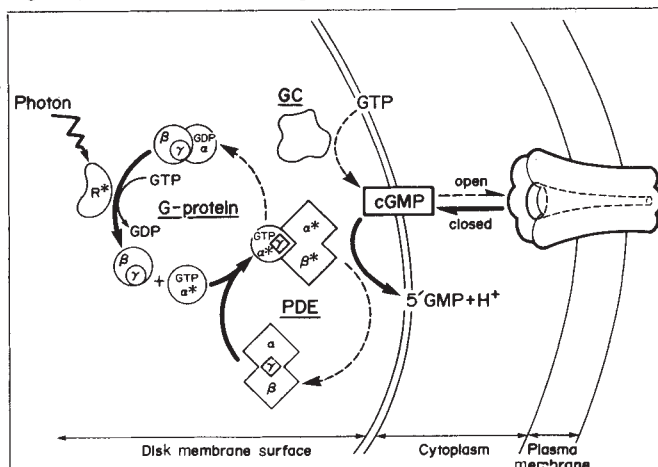
39K channel is effectively blocked.

One of the most surprising results is the relative amounts of each polypeptide. Each group calculates that its molecule is present at about 1–2 per cent rhodopsin (wt/wt). Depending on the subunit composition and molecular mass, this is roughly 1 channel: 500 rhodopsins, about 10-fold greater than predicted if channels are in the plasma membrane only at the density of about 500 channels per μm^2 estimated by patch clamping⁹. Unless there are non-functional channels in the plasma membrane, the excess copies seem likely to be in the disk membranes. Why are the channels so abundant?

Is reconciliation of the two sets of results possible? Both groups use bovine retinas to isolate the channel, but the procedures differ significantly. Cook *et al.*⁴ use CHAPS as a solubilizing detergent, ion-exchange and triazine dye-affinity chromatography and dithiothreitol, a disulphide-reducing agent, throughout the preparation. Matesic and Liebman⁵ use sodium cholate, rely on a tendency of the protein to aggregate in the absence of excess phospholipids, enrich the 'aggregate' by size-exclusion chromatography and use no dithiothreitol. Neither group has achieved absolute purification, but they each make good arguments that the residual contaminants, a 240K polypeptide in the case of Cook *et al.* and rhodopsin in the case of Matesic and Liebman, do not contribute to activity.

The use of dithiothreitol prompts the suggestion that the 63K unit may be a dimer of 39K. As witnessed by properties of rhodopsin, SDS polyacrylamide gel electrophoresis does not always reliably predict the size of membrane proteins; the mobilities of 41K rhodopsin monomer and its dimer are consistent with those of 35K and 60K protein standards. Moreover, dithiothreitol can cause eventual disulphide reduction and anomalous reformation of disulphide bonds. Matesic and Liebman report that they occasionally observe dimers on denaturing gels, but the denaturing gels of Cook *et al.* show no 'monomer' as might be expected if this hypothesis is correct. Alternatively, could the 39K protein be a proteolytically clipped 63K? A more interesting possibility is that the native channel is a heteromeric protein composed of both types of subunits. If so, each type of subunit can reassemble independently to form a cGMP-dependent ion conductance, albeit with slightly different characteristics.

Are there two channels? There may be



Molecular scheme for rod photoreceptor signalling. Light absorbed by the signal-receptor rhodopsin is transduced to yield an activated receptor (R^*). R^* catalyses the exchange of GTP for GDP bound to the α -subunit of G protein (transducin). In the presence of GTP, the α -subunit (α^*) dissociates from the $\beta\gamma$ -subunit and can activate the cGMP phosphodiesterase (PDE) by displacing or removing the small inhibitory subunit (γ) of PDE. This cascade of events is amplified because one R^* can catalyse the activation of several hundred G proteins and each activated PDE can hydrolyse about 1,000 cGMP per second. Thus, cGMP is rapidly depleted. As it drops below 10 μM , bound cGMP dissociates from the cGMP-dependent ion channel in the plasma membrane. The channel closes to produce membrane hyperpolarization — the cell signal that may be transmitted to the brain. Guanylyl cyclase (GC) returns the cGMP to resting concentration. Shut-off and modulation of the activated components in the cGMP cascade are less well understood.

about 1–3 $\text{ms}^{10,11}$. The only well-characterized antagonist¹² of the channel is L-cis-diltiazem.

In the absence of a neurotoxin to tag the light-suppressible channel, cGMP-dependent conductances have been assayed by reconstituting solubilized fractions in lipid vesicles and measuring cGMP-induced Ca^{2+} fluxes. The 63K molecule isolated¹ by Cook *et al.* has higher affinity for cGMP and shows greater cooperativity ($K_{1/2} = 11 \mu\text{M}$, $n = 3.1$) than the 39K protein of Matesic and Liebman reported in this issue⁵, which behaves at the opposite extreme ($K_{1/2} = 62 \mu\text{M}$, $n = 1.7$). The latter protein is photoaffinity-labelled with 8-azido cGMP. Significantly, the reconstituted 63K channel is unaffected by L-cis-diltiazem, whereas the

two cGMP-dependent conductances. Usually, the presence of cone photoreceptor material in a bovine retina preparation is thought to be minimal, but is it possible that one channel is from cones and the other from rods? Cook *et al.* find that L-cis-diltiazem perfused into the rod cell does not block the light-sensitive conductance and suggest that their molecule has properties of this channel form. There may, however, be other reasons, such as hydrolysis or competition, why the perfusion does not work. There is no other compelling physiological evidence that more than one channel exists. If there are two in one cell type, why?

Both groups should be able to resolve whether the polypeptides are related by the combined use of discriminating antibodies, peptide mapping and sequencing. Then we can look forward to learning just how this type of channel is controlled. Is it modulated by phosphorylation or by interaction with other regulatory proteins? Cyclic-nucleotide-gated conductances are

now being identified in other sensory systems and may well reside in other neurons¹³. With the biochemical questions dispelled, this channel promises to serve as an informative molecular model for cyclic-nucleotide-gated ion channels. □

1. Altman, J. *Nature* **313**, 264-265 (1985).
2. Applebury, M.L. & Hargrave, P.A. *Vision Res.* **26**, 1881-1895 (1986).
3. Stryer, L. *A. Rev. Neurosci.* **9**, 87-119 (1986).
4. Cook, N.J., Hanke, W. & Kaupp, U.B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 585-589 (1987).
5. Matesic, D. & Liebman, P.A. *Nature* **326**, 600-603 (1987).
6. Nicol, G.D. & Miller, W.H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5217-5220 (1978).
7. Caretta, A., Cavagioni, A. & Sorbi, R.T. *J. Physiol., Lond.* **295**, 181-178 (1979).
8. Fesenko, E.E., Kolesnikov, S.S. & Lyubarsky, A.L. *Nature* **313**, 310-313 (1985).
9. Pugh, E.N. & Cobbs, W.H. *Vision Res.* **26**, 1613-1643 (1986).
10. Haynes, L.W., Kay, A.R. & Yau, K.-W. *Nature* **321**, 66-70 (1986).
11. Zimmerman, A.L. & Baylor, D.A. *Nature* **321**, 70-72 (1986).
12. Stern, J.H., Kaupp, U.B. & MacLeish, P.R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1163-1167 (1987).
13. *Nature* **325**, 389 (1987).

Meredith L. Applebury is in the Department of Biological Sciences at Purdue University, West Lafayette, Indiana 47907, USA.

Protein structure

Stability of alpha-helices

Thomas E. Creighton

PROTEINS normally exist in well-defined three-dimensional conformations, in contrast to the usual disordered state of other polymers. The physical basis of the stabilities of these folded conformations is only partially understood, even though they are essential for most biological phenomena. Best understood seemed to be the relatively simple isolated α -helix, but recent work by Baldwin and co-workers, reported on page 563 of this issue¹, identifies one factor that has been overlooked and indicates that others remain to be discovered.

The formation of α -helices in disordered polypeptides is a classical nucleation event, usually treated by the Zimm-Bragg theory². Formation of the first segment of the helix from the disordered polypeptide is rate-limiting and energetically unfavourable, with an equilibrium constant, σ , typically of the order of 10^{-4} . Addition of amino-acid residues to either end of the helix is very rapid and occurs with an equilibrium constant, s , close to unity. The unfavourable nature of the nucleation event probably results from the need to fix, simultaneously and independently, at least four contiguous residues in the helical conformation, to give at least one complete helical turn and the first hydrogen bond; in contrast, growth of the helix occurs by addition of individual residues (see figure).

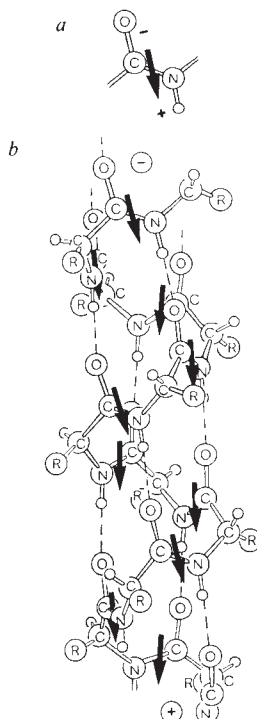
Different types of amino-acid residues have different values of s , but the Zimm-Bragg theory assumes that this single

value pertains to both ends of the helix, and is independent of the neighbouring sequence. Using this theory, and measuring the effect of a few 'guest' amino-acid residues on the helicity of a 'host' polypeptide, Scheraga *et al.* have measured the values of σ and s for 18 of the 20

amino-acid residues. The values of s at 25 °C range from 0.6 for Gly to 1.32 for nonionized Glu³. The Zimm-Bragg theory predicts that only polypeptides of amino acids with $s \geq 1$ should form stable helices, and then only with long polypeptides. With $\sigma = 10^{-4}$, for example, a long homopolypeptide with $s = 1$ would consist of stretches of helix interspersed with disordered segments, each averaging 100 residues in length. Short peptides of less than 20 residues are not expected to form stable helices; for example, even with the maximum observed value of $s = 1.3$, a 13-residue peptide would be expected to be helical only 8 per cent of the time.

The model explains most observations with homopolypeptides and the usual instability of helices in isolated peptides, but there are some exceptions. Ooi and co-workers⁴, for example, observed greater helical content in a segment of Ala residues when in the sequence Glu₂₀-Ala₂₀-Phe than in Ala₂₀-Glu₂₀-Phe, suggesting that the value of s is dependent on position in the polypeptide. Also, some small peptides have much greater than expected helical contents. Brown and Klee⁵, for example, observed significant amounts of helix in the C-peptide, although only at low temperatures. (C- and S-peptides are cleaved from ribonuclease A by cyanogen bromide and subtilisin, respectively.) The C-peptide, consisting of the 13 amino-terminal residues of ribonuclease A, contains no residues with $s > 1.09$; residues 3-13 exist in a stable helical conformation in folded ribonuclease A. Using the C-peptide, the S-peptide (residues 1-20), and synthetic homologues, these observations are now confirmed and extended by Baldwin and co-workers¹ and by Rico *et al.*⁶. These authors observe helical contents of up to about 50 per cent and localize the helix to only the amino-terminal portion of the S-peptide⁷, as in folded ribonuclease. The variation of helicity with pH and with salt concentration indicates that electrostatic interactions are partially responsible, but the initial suggestion of a particular stabilizing salt-bridge between side chains was disproved. Nevertheless, the effects of charged residues are dependent on their position in the peptide. The helical content is increased by positively charged groups near the carboxyl end and by negatively charged groups at the amino end, and is decreased when they are at the other ends. There is a similar pattern in folded proteins, where there is a tendency for charged residues to occur at that end of the helix at which they stabilize it⁸.

The explanation offered for these observations is that the charged groups interact with the dipole of the helix, as had been suggested by Blagdon and Goodman⁸. The helix dipole arises from the substantial dipole of the peptide bond (about 3.5 Debye units, or 1.2×10^{-29} Cm),



a, Dipole of a single peptide bond. b, Individual peptide-bond dipoles within an α -helix, showing how they are aligned nearly head-to-tail.

several of which are aligned nearly head-to-tail in the α -helix. Consequently, each end of the helix may have an effective net charge of about half an electron, being positive at the amino end. Hol has recently proposed⁹ roles for the helix dipole in protein structure and function, particularly for ligand binding. Earlier, Brant and Flory¹⁰ stressed the importance of the helix dipole for protein conformation and helix stability, pointing out that energetically unfavourable parallel alignments of the peptide-bond dipoles would inhibit nucleation of the helix, whereas its further growth would be aided by head-to-tail alignment of the dipoles. The new observations^{1,5} indicate that the helical state is stabilized by interactions of oppositely charged groups with either end of the helix dipole, and that it is destabilized by unfavourable interactions.

These electrostatic interactions, however, have only a relatively small effect; a single charge interaction has been shown to alter the helical content by at most threefold, and usually by only twofold. Even without such stabilizing interactions, the S-peptide α -helix is about 30-fold more stable than would have been predicted, so the major interactions that stabilize this particular helix remain unknown.

Most important, systematic studies of helix stability in small peptides are now feasible, and Scheraga and co-workers³ are using the experimental data to develop an extension of the Zimm-Brugg theory that will take into account interactions of each side chain with the helix dipole and with other side chains. Predicting the helical propensities of different amino-acid sequences should be greatly improved, certainly when in an isolated peptide and possibly within folded proteins. It seems clear, however, that helical conformations in folded proteins are ultimately determined by tertiary interactions with the rest of the protein¹¹. Nevertheless, Mitchinson and Baldwin¹² were able to produce semisynthetic forms of ribonuclease S with increased thermostabilities by combining their more helical S-peptide homologues with the normal S-protein. Stabilizing proteins by stabilizing their α -helices, and perhaps other elements of secondary structure, should become possible whenever the required covalent structure changes do not affect the internal packing or other stabilizing interactions of the folded protein.

Because α -helices are favourite candidates for nucleation sites in folding proteins, the greater than expected stability of some may tend to increase this popularity. But in my opinion, the observed kinetics of protein unfolding and refolding are inconsistent with a nucleation event being the overall rate-limiting step, like that in helix formation. Partially folded intermediates undoubtedly are important before the rate-limiting step, and helices

Variations of AIDS virus relatives

ALTHOUGH one of the hallmarks of HIV-1, the human immunodeficiency virus type 1 that causes AIDS, is the considerable variability between independent isolates, on page 610 of this issue James Mullins and his colleagues report¹ that a remarkable invariability seems to characterize isolates of a related group of viruses with a somewhat nebulous relationship to AIDS. The alternative interpretation considered by the authors — and the subject of rumours for some months — is that the viruses are not all independent isolates but rather the result of laboratory contamination.

The group of viruses studied by Mullins's team comprises three monkey viruses and one human virus. The latter is HTLV-4, the human T-cell lymphotropic virus type 4, which was first isolated from healthy Senegalese and has since been shown by serological means to be present elsewhere in western Africa, including Ivory Coast, where it is widely prevalent². The three monkey viruses are STLV-3_{mac}, STLV-3_{agn} and STLV-3_{smm}, the simian T-cell lymphotropic viruses type 3 isolated from the rhesus macaque, the African green monkey and the sooty mangabey monkey, respectively. Only the first of these seems to cause immunodeficiency in its host.

In a paper published at the tail end of last year Mullins, M. Essex and colleagues described³ the molecular cloning of STLV-3_{agn} from an infected cell line and a consensus restriction-enzyme map of the predominant form of the virus in the cells. Mullins and his collaborators have now been able to compare this map with maps of HTLV-4 (which they cloned from an infected cell line provided by the Essex laboratory), STLV-3_{mac} (provided by R. Desrosiers) and additional isolates of STLV-3_{agn} (provided by Essex). What is remarkable about these maps is that they are all much the same, although a few variations are detected both between the different viruses and between different isolates of the same virus type.

The similarity is all the more remarkable when HIV-2, the virus isolated by Luc Montagnier and his colleagues from AIDS

patients in western Africa, is taken into account. According to Montagnier's group⁴, HIV-2 is prevalent in the same areas as HTLV-4 and the two are clearly closely related. And yet whereas the restriction maps of the three isolates of HTLV-4 reported in this issue are almost identical to each other, that of an isolate of HIV-2 is very different and there is considerable variation among HIV-2 isolates⁴.

Mullins and colleagues suggest two interpretations of these data. The first is that "STLV-3_{agn}, STLV-3_{mac} and HTLV-4 may represent a remarkably stable virus capable of transmission between several species". In favour of that interpretation, they note that two of the STLV-3_{mac} viruses were isolated three years apart and yet have identical restriction maps. They also note that HTLV-1 and STLV-1 are very similar in structure and each can infect other species. The alternative interpretation put forward is that some of the viruses "do not correspond to independent virus isolates but originate from other virus-producing cell cultures maintained in the same laboratories"¹. If that is the case, the most likely possibility is that STLV-3_{mac} is masquerading as STLV-3_{agn} and HTLV-4.

If contamination can be excluded — and that will probably require the isolation under stringent conditions of further examples of the viruses — there will be several very interesting questions to answer. Why is there such stability in one group of viruses but such variability in others? Does it reflect a shift from a non-pathogenic virus, HTLV-4, to a pathogenic form, HIV-2? And, if it does, what differences in sequence account for the shift? Whatever the real identity of the cloned HTLV-4, sight should not be lost of the fact that 1–5 per cent of several populations of west Africans are infected with HIV-2/HTLV-4 and are consequently at some risk of developing AIDS. □

1. Kornfeld, H., Riedel, N., Vighianti, G., Hirsch, V. & Mullins, J.I. *Nature* 326, 610–613 (1987).
2. Denis, F. *et al. Lancet* i, 408–411 (1987).
3. Hirsch, V. *et al. Proc. natn. Acad. Sci. U.S.A.* 83, 9754–9758 (1986).
4. Clavel, F. *et al. Nature* 324, 691–695 (1986).

Peter Newmark is Deputy Editor of Nature.

may play a role in stabilizing them. Yet the surprisingly stable S-peptide helix I have discussed here is reported to appear only at the same time as most of the ribonuclease tertiary structure, not in the observed initial steps of folding¹³. Perhaps too much emphasis has been placed on elements of secondary structure. □

1. Shoemaker, K.R., Kim, P.S., York, E.J., Stewart, J.M. & Baldwin, R.L. *Nature* 326, 563–567 (1987).
2. Zimm, B.H. & Bragg, J.K. *J. chem. Phys.* 31, 526–535 (1959).
3. Vasquez, B.H., Pincus, M.R. & Scheraga, H.A. *Biopolymers* 26, 351–371 (1987).
4. Ihara, S., Ooi, T. & Takahashi, S. *Biopolymers* 21, 131–145 (1982).

5. Brown, J.E. & Klee, W.A. *Biochemistry* 10, 470–476 (1971).
6. Rico, M. *et al. Biopolymers* 25, 1031–1053 (1986).
7. Kim, P.S. & Baldwin, R.L. *Nature* 307, 329–334 (1984).
8. Blagdon, D.E. & Goodman, M. *Biopolymers* 14, 241–245 (1975).
9. Hol, W.G.J. *Prog. Biophys. molec. Biol.* 45, 149–195 (1985).
10. Brant, D.A. & Flory, P.J. *J. Am. chem. Soc.* 87, 2791–2800 (1965).
11. Kabsch, W. & Sander, C. *Proc. natn. Acad. Sci. U.S.A.* 81, 1075–1078 (1984).
12. Mitchinson, C. & Baldwin, R.L. *Prot. Struct. Funct. Genet.* 1, 23–33 (1986).
13. Brems, D.N. & Baldwin, R.L. *J. molec. Biol.* 18, 1141–1156 (1984).

Thomas E. Creighton is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Insulin, IGF-I and growth in diabetic rats

SIR—Scheiwiller *et al.* recently presented evidence¹ that insulin, and to a lesser extent IGF-I (insulin-like growth factor I), stimulate growth in diabetic rats, and proposed "a new mechanism of endocrine control of growth", in which growth hormone (GH) and insulin act mainly by modulating hepatic IGF-I synthesis.

It is of course well-known that diabetic animals fail to thrive and grow, an impairment which is restored by insulin treatment². Scheiwiller *et al.* make no mention of an obvious explanation for their results, namely that diabetes obliterates pulsatile GH secretion, whereas insulin treatment restores GH secretion and with it, growth. Diabetic rats show greatly impaired GH secretion³. We have found that insulin treatment does restore GH secretory patterns to normal in diabetic rats, and restores growth (see figure).

Scheiwiller *et al.* suggested that the growth-promoting and hypoglycaemic effects of insulin and IGF-I were independent because hyperglycaemia persisted in the animals receiving all but the highest doses of insulin. In our experiments, the return of GH pulsatility and reductions in food intake and urine output occurred soon after insulin treatment, but well before blood sugar values fully returned to normal. Indeed, diabetic animals receiving insulin gained weight immediately owing to the early amelioration of diabetic symptoms, followed by a much more gradual increase reflecting an anabolic response. It would be interesting to learn what proportion of the 6-day body-weight increases reported by Scheiwiller *et al.* occurred within the first two days of insulin or IGF-I treatment. The same group has also found that insulin is an order of magnitude more potent as a growth promoter than IGF-I, by a factor which closely parallels the relative hypoglycaemic potencies of these peptides⁴. We agree that growth can proceed before euglycaemia is reached, so the relevant parameter may be the relative abilities of IGF-I and insulin to restore intracellular glucose transport in diabetic rats.

We propose a different explanation. Intracellular starvation blocks anabolic responses in the absence of an insulin-like substance, GH secretion is suppressed and growth is arrested. Without normal GH secretion, IGF-I levels fall. Treat-

ment with insulin (or larger doses of less potent insulin-like substances) corrects intracellular metabolism, restores the ability to produce an anabolic response, restores normal GH secretion and allows growth to resume. The resumption of GH secretion raises IGF-I levels to normal. We suggest it is unnecessary to advocate new endocrine mechanisms controlling growth without properly considering the most potent old one — endogenous growth hormone itself.

I.C.A.F. ROBINSON

R.G. CLARK

L.M.S. CARLSSON

National Institute for Medical Research,
Mill Hill, London NW7 1AA, UK

1. Scheiwiller, E. *et al.* *Nature* **323**, 169–171 (1986).
2. Fajans, S.S. in *Endocrinology* Vol. 2 (ed. DeGroot, L.J.), 1007–1023 (Grune and Stratton, New York, 1979).
3. Tannenbaum, G.S. *Endocrinology* **108**, 76–82 (1981).
4. Froesch, E.R. *et al.* *J. Endocr.* **111**, (suppl.), 2P (1986).
5. Clark, R. G. *et al.* *J. Endocr.* **111**, 27–35 (1986).

FROESCH AND ZAPF REPLY—Robinson, Clark and Carlsson demonstrate very beautifully that the restoration of normal glycaemia in diabetic rats by insulin normalizes the secretory pattern of growth hormone. However, we do not agree with their view that the effects of IGF-I on growth simply reflect "insulin-like" potency and a correction of intracellular metabolism. It is accepted that the relative 'growth potency' of IGF-I differs from that of insulin by a factor of 100 or more depending on the cell type¹. IGF-I in very small concentrations affects differentiation and replication of chondrocytes and calvaria cells^{2–4} but much larger concentrations are required for acute insulin-like effects on muscle and adipocytes⁵.

We would like to add the relevant data on the severity of diabetes in the different groups of rats. Glucosuria was identical in untreated diabetic rats and in growing diabetic rats treated with IGF-I (300 µg day⁻¹) and in both groups averaged 30 g per 24 h per 100 g body weight. In the normoglycaemic insulin-treated (2.5 U per day) rats, glucosuria was reduced to 3 g per day. Body weight gain was linear over 6 days in the insulin- and IGF-I-treated groups. The severity of diabetes is also reflected by the food consumption which averaged (in g per day): 21 in untreated diabetic rats; 7 in control rats; 24 in diabetic rats treated with 300 µg of IGF 1 per day; and 13 in diabetic, insulin-treated normoglycaemic rats, showing that the diabetic state was nearly normalized by

insulin but not by recombinant human (rh) IGF-I. The growth-promoting effect of either insulin or rhIGF-I does not therefore parallel the overall anti-diabetic and blood glucose-lowering potency. We conclude that IGF-treated diabetic rats are almost as diabetic as untreated diabetic rats but grow considerably despite hyperglycaemia; that the degree of stimulation of growth parameters correlates very well with serum IGF-I levels, and that IGF-I can stimulate growth of hypophysectomized rats without lowering blood sugar^{5,6}.

It seems to us that the question is not whether IGF-I by itself has growth-promoting activity but rather if growth hormone and IGF-I can act on growth in a concerted way. It would be interesting to infuse diabetic rats with IGF-I to find out whether IGF-I induces rat growth hormone peaks in a similar way to insulin.

E.R. FROESCH

J. ZAPF

Metabolic Unit, Department of Medicine,
University of Zurich,
CH 8091 Zurich, Switzerland

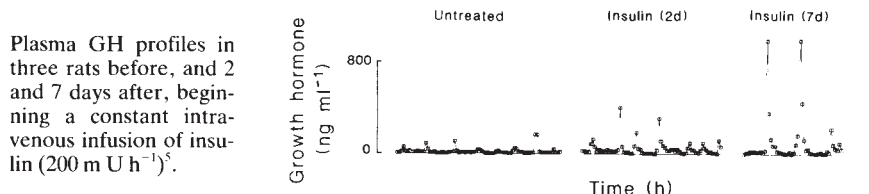
1. Froesch E.R. & Zapf, J. *Diabetologia*, **28**, 485–493 (1985).
2. Schmid, Ch., Steiner, Th. & Froesch, E.R. *Calcif. Tissue Int.* **35**, 578–585 (1983).
3. Schmid, Ch., Steiner, Th. & Froesch, E.R. *FEBS Lett.* **173**, 48–52 (1984).
4. Vetter, U. *et al.* *J. clin. Invest.* **77**, 1903–1908 (1986).
5. Schoenle, E., Zapf, J., Humbel, R.E. & Froesch, E.R. *Nature* **296**, 252–253 (1982).
6. Schoenle, E., Zapf, J., Hauri, Ch., Steiner, Th. & Froesch, E.R. *Acta endocr.* **108**, 167–174 (1985).

AIDS vaccine predictions

SIR—Although most of the attempts to develop vaccines against AIDS (acquired immune deficiency syndrome) have focused on the B-lymphocyte defined epitopes of the envelope (env) protein of the human immunodeficiency virus (HIV-1) (see refs 1 and 2) there is a growing interest in other proteins of the virus because, for instance, patients with AIDS can die with high levels of antibody to the env protein in their blood³. Greater emphasis is therefore starting to be placed on both other proteins and the role of cell-mediated immunity, particularly cytotoxic T-cells, which may provide resistance to HIV-1. T cells recognize parts of proteins (T epitopes) which are often different from B cell epitopes^{4,5}. Thus the best candidate for a peptide vaccine might be a chain section containing both B and T epitopes in regions conserved in sequence between the different isolates.

The immune response against the proteins of the gag gene might be crucial, as clinical progression to AIDS has been associated with a reduction in antibodies to gag p24, the main capsid protein⁶. Thus a gag p24 vaccine might prevent the onset of AIDS in HIV-1 carriers. In addition, a prophylactic vaccine might require both env and gag components.

Therefore, as shown in the figure, we



1. Pologe, L.G. & Ravetch, J.V. *Nature* **322**, 474-477 (1986).
2. Pologe, L.G., Pavlovic, A., Shio, H. & Ravetch, J.V. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Great stinking fogs

Eric Ashby

The Big Smoke: A History of Air Pollution in London Since Medieval Times.

By Peter Brimblecombe. Methuen: 1987. Pp. 185. £25, \$39.95.

Air pollution rises and falls among the priorities of public anxiety. After events such as the 1952 London smog, or Chernobyl, it shoots up to the top of the charts. Then, after a few months, the media get tired of the issue and the public turns its attention back to the perennial issues that plague affluent societies: violent crime, drugs, unemployment, budget-deficits and — for the next few years at least — AIDS.

But among people whose business it is to think about pollution there is a sustained concern about the matter. It is only the agenda that changes: at one time smoke, at another sulphur dioxide, then ozone, lead, radioactivity, chlorofluorocarbons. Acid rain remains high on the agenda, and now that an 'ozone-hole' has been found in the Antarctic (*Nature* 324, 651; 1986), organic chlorine compounds are back in the news. Then, of course, there is the global rise in levels of carbon dioxide and the consequent possibility of a 'greenhouse effect', which a US Senate Committee regards as "no longer a question of science [but] now one of policy". While scientists and politicians busy themselves with what should be done about this and other problems of pollution, it is important that some people should put the issues into the framework of history. That is what Peter Brimblecombe has done in this entertaining and scholarly book.

The history of smoke pollution in England since the seventeenth century has been well documented, but most of the material is in reports of commissions and committees which even the earnest student finds dreary to read. There have been a few books which summarize the documents and interpret them for the general reader, but their scope has been limited — either they have gone back in time only to the industrial revolution, with, perhaps, a passing wave-of-the-hand to John Evelyn's *Fumifugium*, or they have concentrated upon the technology or the politics of smoke-control. The merit of *The Big Smoke* is that it is free from these constraints. Brimblecombe goes back to the Bronze Age and has gathered evidence not only from technology and political science but also from art, drama and literature.

The book opens with a vivid example. When Britain was a Roman colony smoke was already a hazard to health, for people kept themselves warm with fires in huts

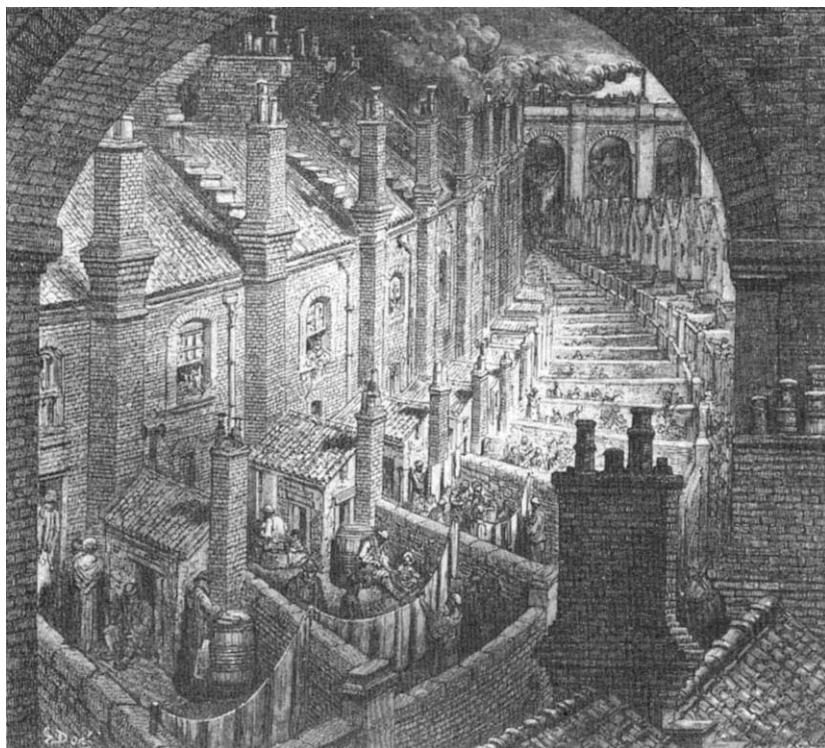
without chimneys. There is circumstantial evidence that this indoor pollution affected their health; it comes from an examination of thousands of skulls from Bronze Age to mediaeval times. From pitting in the bones of Anglo-Saxon skulls it seems that sinusitis was common, and bad ventilation is a contributory cause of the complaint. Coupled with this is evidence of blackening of the lungs (anthracosis) in mummified tissues. Chimneys, says Brimblecombe, were probably not common in any but noble households much before the late sixteenth century.

In mediaeval times timber was the main source of fuel (though blacksmiths' forges and lime kilns used coal, and there are records of litigation when culprits were brought to court for fouling the air). Then came London's first fuel 'crisis' when the forests within easy distance of the city had been denuded. The fuel for the future became coal, brought by sea from Newcastle; and with coal came the "Hellish and dismall Cloud" over London described by John Evelyn. Between 1580 and 1680 imports of coal into London grew from some 20,000 tons a year to over 360,000 tons. With this growth came

"Great Stinking Fogs", decayed stonework and ill health for the inhabitants.

Brimblecombe describes the effects on London's citizens. It is a picture only too familiar to us in the 1980s: a beneficial technology (pesticides, nuclear power) that has unanticipated and disturbing side-effects. The befogging of London fortunately coincided with a fresh intellectual clarity, for the seventeenth century was the age of the philosopher-scientists and their meetings as Fellows of the newly founded Royal Society were often devoted to discussions about smoke abatement. One of them, John Graunt, published the first convincing statistics of mortality due to fogs. Another, John Evelyn, examined the possibilities of using coke, from which the obnoxious gases had been removed; he also proposed that industries should be moved out of London into what we would now call industrial estates. And Henri Justel published in the Society's *Transactions* the description of a stove which, he claimed, would consume its own smoke, and was so successful that "coal steeped in Cats-piss makes not the least ill scent".

The book then carries the reader into the nineteenth century, with a summary of a familiar history — the efforts of technologists to abate smoke and of politicians to legislate for clean air; of reformers working in the public interest to cut a path through jungles of resistance; of the selfishness of mill owners, the myopia of politicians, the inertia of civil servants and the sentimentality of householders who



Urban grime — how the slums of London appeared in 1870, as engraved by Gustave Doré.

regarded the open coal fire in the living room as a symbol of home, not to be exchanged for a gas fire or a closed anthracite stove. It is a task not yet over: the Control of Pollution Act 1974 has not yet been fully implemented. If there is a lesson in Peter Brimblecombe's story it is about the century-long delay between the adoption of a technological change and the adaptation of society to cope with the change.

But the refreshing thing about *The Big Smoke* is that it is not written to teach lessons; it is out to tell a story. Brimblecombe writes in a relaxed, 'effortless' style (I put the word in quotation marks because of course it was not effortless: it is the product of careful craftsmanship). The book is full of fascinating digressions. Thus most people know of the polluter-pays-principle, but not that it was introduced by Archbishop Laud over 300 years ago, when he obliged the brewers in Westminster to pay dues for the use of coal, which he then used to defray the cost of repairing pollution-damage to St Paul's Cathedral. Long-range transport of pollutants in the air is high on the agenda of the 1980s, but plumes of soot from London were observed 50 miles away by the naturalist Gilbert White 200 years ago. Anyone old enough to recollect pea-soup fogs knows how quickly clothes became dirty, wallpaper discoloured and curtains shabby; Brimblecombe gives glimpses of the effect this had on fashions and furnishings, an effect that still persists in the sombre rooms of some London clubs.

Most entertaining of the author's embellishments to his theme are the references to fog and smoke in literature. That Sherlock Holmes operated in fogs is well known; so are the fogs in Dickens's London. But it needs a sensitive scholar to discover what people thought about the murky climate, in poems by Swift and Gay, in essays by Lamb (who talked about the city's "beloved smoke") and in D.H. Lawrence's *Lady Chatterley's Lover* (where he indicts coal as "excrement" and its pollution as "black manna from the skies of doom"). Artists, too, have recorded the social response to smoke, and *The Big Smoke* is enriched by some ominous engravings by Gustave Doré. Indeed, these and dozens of other illustrations greatly increase the interest of the book.

I read this book with admiration, enjoyment and (I confess) a touch of envy. It does not have the authority of a masterpiece such as Braudel's *Civilization and Capitalism*, but I detected in it echoes of that master's informal and intimate style. □

Lord Ashby is a Fellow of Clare College, Cambridge CB2 1TL, UK. He was the first chairman of the Royal Commission on Environmental Pollution, and co-author with Mary Anderson of The Politics of Clean Air (Oxford University Press, 1981).

Antibodies in the laboratory

Richard Batchelor & colleagues

Handbook of Experimental Immunology, 4th Edn. Edited by D.M. Weir *et al.* Blackwell Scientific: 1986. Vol. 1. pp. 876, £75, \$95; Vol. 2 pp. 624, £48, \$95; Vol. 3 pp. 324, £25, \$80; Vol. 4 pp. 568, £45, \$80. Four-volume set £180, \$315.

ON LEARNING that a new edition of this work has been published, the first reaction of immunologists should properly be one of gratitude to the editors for taking on such an ambitious task. How well have they succeeded with this fourth edition?

The increased size of the *Handbook*, as the senior editor comments, reflects the expansion of immunology itself. Not only is there now an extra volume devoted to genetics and molecular immunology, but the number of chapters in the other volumes has also increased; indeed, the list of contributors occupies nearly five pages. One new feature is that many sections are introduced with 'overviews' in an attempt to place the different topics in context. The remaining contributions provide laboratory protocols, with accompanying theoretical explanations and notes to alert users to special difficulties.

Volume 1 (*Immunochemistry*) now contains 40 chapters. Some of these are excellent, in particular "Preparation and Purification of Active Fragments from Mouse Monoclonal Antibodies", "Glycoprotein Antigens of the Lymphocyte Surface and Their Purification by Affinity Chromatography", "Kinetics of Antibody Reactions and the Analysis of Cell Surface Antigens" and "Complement Technology", all of which will be of use to both those with and those without some experience in the area. Sadly, though, some of the other contributions are disappointing. There are also omissions: why, when there is a good chapter on the identification of antigens and antibodies by electron microscopy, is there no contribution on immunocytochemistry at the light microscopy level? Similarly, there is a detailed chapter on biochemical preparative and analytical methods applied to class I MHC antigens, but no equivalent account for the class II series.

Volume 2 (*Cellular Immunology*) has sections on phagocytes, the lymphoid system, lymphocyte responses and immunoregulation. It begins with two clear overviews of receptors of phagocytes and the physiology of the mononuclear phagocyte system. Subsequently, procedures for studying macrophage markers, phagocytosis and intracellular killing of bacteria, macrophage secretion products, dendritic cells and many other

topics are described. In general these contributions are clearly presented and the choice of subjects will satisfy most bench workers. There are occasional oddities; for example, in Chapter 44 we are instructed on how to count cells with an electronic particle counter but are told nothing about the use of a counting chamber and microscope.

The section on the lymphoid system and lymphocyte responses is the main dish in this volume, and many helpful chapters are included — those dealing with the preparation of lymphocyte subpopulations, genetic markers for following cell populations, and assays for Ig-secreting cells and cytotoxic cells, for example. There are, however, serious deficiencies. Lymphokines are given inadequate space, and anyone needing details on assays or monoclonal antibodies will not find them here.

Immunoregulation is a particularly challenging topic — almost a skeleton in the immunologists' cupboard — and one which could do with less theorizing and more firm experimental data. A little honest admission of ignorance would frequently not come amiss. There are several overviews dealing with T suppressor cells, and it must be questioned whether the editors have achieved a satisfactory balance here. Thus every one of the 13 references in the contribution on epitope-specific suppression is to work from the author's own laboratory. Nowhere is there any satisfactory discussion of the really awkward issues, such as the nature of what is still improperly called I-J, and whether specific suppressors are distinct from helper and cytotoxic T cells.

Volume 3 (*Genetics and Molecular Immunology*) is largely a series of reviews, mixing theoretical discussion with some useful reference tables on such topics as murine recombinant MHC haplotypes. Most of the authors make no attempt to provide a methods manual. There are two exceptions: the chapter on genomic gene cloning in bacteriophage lambda and that on gene transfection for lymphocyte surface antigens. This volume is heavily biased towards immunoglobulin genetics, as declared in the first chapter: "Molecular Immunology is, today, largely the study of the nucleic acids involved in the specification of immunoglobulin structure and concentration". This may have been true five years ago, but not now. Among the omissions here are T-cell receptor and lymphokine genetics.

The reviews of IgH genes are clear and well-presented; but to have seven chapters devoted to Ig allotypes seems inappropriate, particularly in light of the omissions. For example, there is no discussion of the mechanisms of somatic mutation or of allelic exclusion.

The other main topic covered in this volume is the MHC. The chapters cover-

ing the mouse and rat are excellent and include useful listings of recombinant haplotypes and MHC mutant strains. However that on HLA genetics and disease associations is inadequate — there is no detailed presentation of the organization of the class II (HLA-D) region, and the discussion of the mechanisms of disease association is sketchy.

By the very nature of its intended scope, Vol. 4 (*Applications of Immunological Methods in Biomedical Sciences*) suffers particularly from trying to be all things to all scientists. To select just seven uses of monoclonal antibodies is to present only the tip of the iceberg. Although a good chapter on the preparation of immunotoxins appears in Vol. 1, there is no counterpart here on their clinical use. The chapter on neurobiology is excellent — but who would have expected to find it in this volume? And what of all the other cell systems in which monoclonal antibodies are proving to be invaluable tools? The

very success of immunological techniques, and of monoclonal antibodies in particular, leads one to conclude that the editors set themselves an impossible task with this final volume.

If the aim in revising the *Handbook* was to give a comprehensive theoretical background to immunology, this new edition cannot be counted a success. There are too many omissions and idiosyncratic choices of material. Further, the protocols and references in many chapters are unnecessarily dated. For the future, we would humbly suggest that these defects could be minimized by more rigorous editorial control, staggering of revision of the individual volumes, and use of word-processing and computerized printing methods to reduce the interval between writing and reading. □

Richard Batchelor, Mary Ritter and Robert Lechler are in the Department of Immunology, Royal Postgraduate Medical School, Hammersmith Hospital, London W12 0HS, UK.

Filling the plane

Roger Penrose

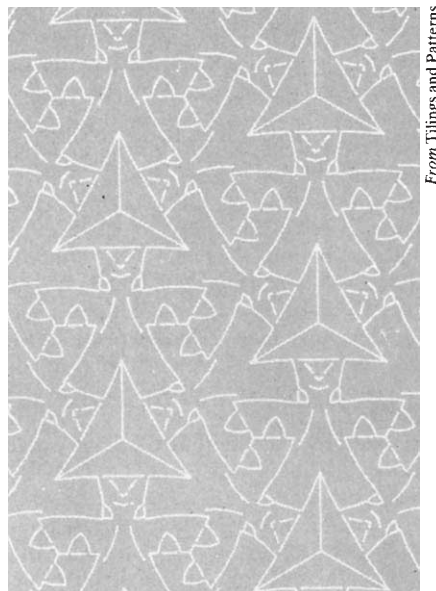
Tilings And Patterns. By Branko Grünbaum and G. C. Shephard. W. H. Freeman: 1987. Pp. 700. \$59.95, £54.95.

A TILING is a covering of the euclidean plane with a (countably infinite) number of shapes whose interiors do not overlap. This is a sort of infinite jig-saw puzzle. A pattern is a somewhat less precise, but related notion, where the plane is marked to illustrate some ordered arrangement, such as occurs with crystalline structures. Instances of tilings are the familiar regular subdivisions of the plane into squares, or equilateral triangles or regular hexagons.

This appears to be a simple enough idea, and it is surprising that before the publication of the book under review no thoroughly systematic account of plane tilings had been put forward. Mathematicians seem to have generally regarded tiling problems for the plane to be essentially trivial. Even the great David Hilbert, in his famous address to the International Congress of Mathematicians in 1900, apparently took this view, proposing as one of his famous problems a question concerning three-dimensional tilings; this, as Heesch showed in 1935, can be resolved (in the negative) with some comparatively simple-looking plane tiles. Moreover, it seems that Grünbaum and Shephard did not themselves immediately see the richness of their subject matter. It is my understanding that their original idea was to compile a systematic account of three-dimensional space-filling tilings and that, as an introduction, they had proposed to write an opening chapter on plane tilings.

That introductory chapter has now grown to a beautifully produced book of 700 pages.

Geometrical tilings and patterns are very ancient, and are to be found in the mosaics of Roman times and earlier. Their classification according to the 17 different possible plane crystallographic symmetry groups is something that has been discussed many times by various authors. A thorough account of this type of classification also appears in the present work, but its main thrust is really of a different character altogether. Symmetry is only one aspect of a plane tiling or pattern, and there are many other properties that cannot be captured from considerations of symmetry alone. The reader is introduced



Human jig-saw — isohedral tiling where the basic tile is in the form of a cloaked figure sporting a triangular hat, a pattern inspired by an Escher sketch.

to many unfamiliar terms used in the study of tilings, such as 'prototile', 'isohedral', 'monohedral', 'isogonal', 'isotaxal', 'n-homeohedral', 'homeomeric' and 'poly-hex'. The authors present many general results, some of which appear here for the first time. The discussions are thorough, systematic, and rigorous, yet with no sacrifice of readability. *Tilings and Patterns* provides an absorbing account of its subject, historical as well as mathematical, and is much enhanced by the large number of fascinating and meticulously executed illustrations.

Especially noteworthy is the substantial chapter on aperiodic tilings, which is where much of the previously unpublished material is to be found. The book's appearance is particularly timely because such tilings have come to prominence recently on account of their close relation to the newly discovered quasi-crystals — apparently crystalline structures which exhibit impossible symmetries for something exactly periodic. (The first announcement of such a structure — an icosahedral phase of aluminium-manganese alloy — was made by Shechtman and his colleagues late in 1984.) The book provides the first detailed account of aperiodic tilings of the plane (but not of euclidean three-space — as would be needed for a complete discussion relevant to quasi-crystals). The set of six tiles found by Raphael Robinson in 1971 is covered in some detail, and there is an extensive description of various sets that I myself found in 1973 and 1974. Further, some ingenious sets of tiles found by Robert Ammann in 1977 are discussed and published in this book for the first time. (I noticed a small error here, in Fig. 10.4.12 (b) on p. 556 — some wrong markings.)

In the following chapter, Wang tilings are described (these being the precursor of Robinson's tiles), in terms of which Robert Berger presented the first aperiodic set — involving over 20,000 tiles! The relation between Wang tiles and Turing machines, which in effect forms the basis of Berger's proof that there is no general algorithm for deciding whether or not a given set of tiles will actually tile the plane, is also considered. This result shows that, despite the apparently childish ingredients, the problems of plane tilings are a genuinely difficult area of mathematics. There must indeed be an unending variety of different ways of tiling the plane, and a complete classification must therefore be impossible!

Despite the impossibility of their task, Grünbaum and Shephard's book is a remarkable achievement. It will surely remain the unique reference work in this area for many years to come. □

Roger Penrose is Rouse Ball Professor of Mathematics in the Mathematical Institute, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, UK.

Ichthyologists all at sea

Alwyne Wheeler

Fishes of the North-eastern Atlantic and the Mediterranean. Edited by P. J. P. Whitehead, M.-L. Bauchot, J.-C. Hureau, J. Nielsen and E. Tortonese. *UNESCO: 1985–1987. Three volumes, pp. 1,473.* * **Smiths' Sea Fishes.** Edited by M. M. Smith and P. C. Heemstra. *Macmillan, South Africa/Springer-Verlag: 1986. Pp. 1,047. R145, DM198.*

THERE are estimated to be about 22,000 different species of fishes living today, more than those of all the other vertebrate groups put together. It is not surprising, then, that books attempting to catalogue all the fishes of a large area are now often multi-authored (and sometimes multi-volumed). Each of the works reviewed here is the result of the efforts of around 70 authors, whose contributions have been edited to a consistent style in one case by a team of five editors, in the other by two.

Fishes of the North-eastern Atlantic and the Mediterranean is a set of three volumes designed for identification of all sea fishes in the area. It has keys for the identification of families (Vol. 1), and to genera within each family, and species within the genera; these are scattered through the text. Each species is illustrated and tersely described, notes are given on habitat, food and reproduction, and the distribution is mapped as well as described. The illustrations (all line drawings) are mostly reproductions of existing artwork. This is a concise but comprehensive account of the fishes of European seas, and of the deep sea to the west of Europe, which will be a standard source of reference for many years.

Smiths' Sea Fishes derives from the original *Sea Fishes of Southern Africa* by the late J. L. B. Smith. It is an impressive volume, co-edited by Smith's widow and P. C. Heemstra, and contains accounts of about 2,150 species which live in the seas around southern Africa. The treatment is largely taxonomic, with keys to identification and a short diagnosis of each species. Many species are illustrated with line drawings, others with monochrome photographs some of which, taken from earlier reproductions, are of poor quality. There are 144 colour plates comprising a number of drawings (retained from earlier

editions of Smith's work) or new photographs. The colour illustrations occupy a discrete section of the book, making reference to them so much easier than in *Sea Fishes of Southern Africa* in which they were placed at random throughout the text.

Comparison with the original work of J. L. B. Smith is inevitable, although unjust in some ways because this is essentially a new book. By inviting an international team of authors to work on it, the editors have to a large extent avoided the elements of parochialism that were previously evident. There are still glimpses of the original however. For instance, the short essay on the sense of pain in fishes is a direct quotation and is an example of

opinion promoted as fact. Thirty years after its first publication, there is considerable evidence for the presence of 'pain blockers' in fishes, as in higher vertebrates, to which it would have been reasonable to refer.

On its appearance in 1949 'J. L. B. Smith' was a trend-setting book, and down the years it has been a vital work of reference. The new *Smiths' Sea Fishes* will undoubtedly serve the same role for the next generation (or two) of ichthyologists. Its coverage is essentially the sea fishes of Africa south of the Equator, but it will be valuable far beyond that area. □

Alwyne Wheeler is in the Department of Zoology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK.

Conducted tours

P. N. Butcher

Conduction in Non-Crystalline Materials. By Sir Nevill Mott. *Clarendon: 1987. Pp. 128. £17.50, \$22.36.*

'SMALL is beautiful' is the maxim that Professor Mott has adopted in his new book. His aim is threefold. First, to provide an introduction for students of both experiment and theory. Secondly, to update the earlier and larger book written in collaboration with E. A. Davis (*Electronic Processes in Non-Crystalline Materials*, a second edition of which was published by Oxford University Press in 1979). Thirdly, to show how the ideas developed in studies of amorphous semiconductors can be applied to vitreous silica, amorphous metals, impurity bands in doped crystalline semiconductors and so on.

The result is a superb read which will be of great interest to anyone who has been associated with the developments in this rapidly expanding subject. The book is illuminated throughout by Professor Mott's physical insight into the materials and the processes that go on in them. Many questions are left unanswered because there are no answers. Therein lies the fascination — one is left thinking about what steps might be taken either in the laboratory or in calculations which could begin to lead towards a resolution of the many mysteries remaining in the behaviour exhibited by disordered conductors. This is the sort of intellectual stimulation that one has come to expect from Professor Mott, and he does not disappoint us.

Readers, however, will be screaming with frustration as one subject after another is brilliantly opened up only to be closed down so as to keep the book within its all-too-brief compass of 113 pages of text. There are copious references, but the experts will have read them already, while

students will rapidly lose themselves in a morass of conflicting views if they follow Professor Mott's directions to the primary literature.

There is a central difficulty which has to be faced in writing about conduction in non-crystalline materials. Even after two decades or more of vigorous research, no generally accepted quantitative model has emerged for any particular material. Tremendous progress has been made with the application of amorphous silicon to devices, following the discovery by the Dundee group that this semiconductor can be doped. There have also been great advances in our understanding of many of the details of electron transport in disordered materials. However, there has been little movement on the fundamental questions. What is the nature of the disorder in any particular material? How should it be described? How does the description of the disorder lead to an understanding of the physics of electron transport in that material? This is precisely why the subject continues to interest so many physicists. It is also why it is so difficult to describe in a textbook, particularly a short one.

Professor Mott reviews a great many questions. He provides some of the answers and points out where serious problems remain. Everyone interested in electron transport should acquire this book; students should have a copy of Mott and Davis to accompany it. □

P. N. Butcher is a Professor in the Department of Physics, University of Warwick, Coventry CV4 7AL, UK.

New in paperback

- *Light and Photosynthesis in Aquatic Ecosystems* by John T. O. Kirk. Publisher is Cambridge University Press. Price is £15, \$24.95. For review see *Nature* 309, 302 (1984).
- *Spinors and Space-Time, Vol. 1: Two-Spinor Calculus and Relativistic Fields* by R. Penrose and W. Rindler. Publisher is Cambridge University Press. Price is £17.50, \$29.95. For review see *Nature* 313, 607 (1985).

* In France available from Librairie de l'Unesco, 7 Place de Fontenoy, 75700 Paris; in Britain from HMSO, PO Box 276, London SW8 5DT; in the United States from Unipub, 4611F Assembly Drive, Lanham, Maryland 20706. Price of each volume is, respectively, FF 250, £37.50 and \$45.

Physics and the riddle of life

M. F. Perutz

Erwin Schrödinger's book What is Life?, published in 1944, drew several of the brightest physicists into molecular biology. But the book's chief merit lies in its rescue from obscurity and popularization of an earlier paper by Timoféeff, Zimmer and Delbrück.

IN THE early 1940s, Schrödinger worked at the Institute for Advanced Studies in Dublin. One day he met P.P. Ewald, another German theoretician who was then Professor at the University of Belfast. Ewald, who had been a student in Göttingen before the First World War, gave Schrödinger a paper that had been published in an obscure journal in 1935¹. It was by N.W. Timoféeff-Ressovsky, K.G. Zimmer and Max Delbrück and was entitled "The Nature of Genetic Mutations and the Structure of the Gene"². Apparently Schrödinger had been interested in that subject for some time, but the paper fascinated him so much that he made it the basis of a series of lectures at Trinity College, Dublin, in February 1943; they were published by Cambridge University Press in the following year under the title *What is Life? The Physical Aspect of the Living Cell*.

The book is written in an engaging, lively, almost poetic style ("The probable life time of a radioactive atom is less predictable than that of a healthy sparrow"). It aroused wide interest, especially among young physicists. Up to 1948 it drew 65 reviews and has probably by now sold 100,000 copies. It is still in print and has become a classic, providing nourishment for historians, sociologists and philosophers of science who have commented on it, on the comments on it or on the comments on the comments on it. A PhD thesis published on the subject in 1979 contains over 120 references, excluding the 65 reviews³. François Jacob has explained the reasons for the book's impact best:

After the war, many young physicists were disgusted by the military use that had been made of atomic energy. Moreover, some of them had wearied of the turn experimental physics had taken . . . of the complexity imposed by the use of big machines. They saw in it the end of a science and looked around for other activities. Some looked to biology with a mixture of diffidence and hope. Diffidence because they had about living beings only the vague notions of the zoology and botany they remembered from school. Hope, because the most famous of their elders had painted biology as full of promise. Niels Bohr saw it as the source of new laws of physics. So did Schrödinger, who foretold revival and exaltation to those entering biology, especially the domain of genetics. To hear one of the fathers of quantum mechanics ask himself: "What is Life?" and to describe heredity in terms of molecular structure, of interatomic

bonds, of thermodynamic stability, sufficed to draw towards biology the enthusiasm of young physicists and to confer on them a certain legitimacy. Their ambitions and their interests were confined to a single problem: the physical nature of the genetic information³.

Seymour Benzer, Maurice Wilkins and Gunther Stent have said that the book was decisive in drawing them from physics into biology. Francis Crick told me that he found it interesting, but that he would have switched to biology anyway. I was



Erwin Schrödinger: engaging writer.

already in the thick of trying to solve the structure of proteins when it was published, and I may have been encouraged by its quotation of Darlington's view that genes are made of protein. Crick wrote in 1965:

On those who came into the subject just after the 1939–1945 war Schrödinger's little book . . . seems to have been peculiarly influential. Its main point — that biology needs the stability of chemical bonds and that only quantum mechanics can explain this — was one that only a physicist would feel it necessary to make, but the book was extremely well written and conveyed in an exciting way the idea that, in biology, molecular explanations would not only be extremely important but also that they were just around the corner. This had been said before, but Schrödinger's book was very timely and attracted people who might otherwise not have entered biology at all⁴.

However, in 1970 Crick added:

I cannot recall any occasion when Jim Watson and I discussed the limitations of Schrödinger's book. I think the main reason for this is that we were strongly influenced by Pauling, who had essentially the correct set of ideas. We therefore never wasted any time discussing whether we should think in the way Schrödinger did or the way Pauling did. It seemed quite obvious to us we should follow Pauling⁵.

Nor can I recall Crick, Watson, Kendrew and myself ever discussing the bearing of Schrödinger's book on structural molecular biology during our years together at the Cavendish Laboratory; Stanley Cohen writes that few of the many scientists participating in Delbrück's phage course at Cold Spring Harbor in 1944 had read Schrödinger "and in all the social and intellectual activities of these [post-war] summers I do not recall any mention of Schrödinger". The participants included such pioneers of molecular genetics and biochemistry as Luria, Hershey, Lwoff, Monod and Brachet. Hence the book does not appear to have had much impact on the people already working in the field.

Schrödinger's book is written for the layman and begins with a chapter entitled "The Classical Physicist's Approach to the Subject". He asks how events in space and time taking place in a living organism can be accounted for by physics and chemistry.

Enough is known about the material structure of life to tell exactly why present-day physics cannot account for life. That difference lies in the statistical point of view. It is well-nigh unthinkable that the laws and regulations thus discovered [i.e. by physics] should apply immediately to the behaviour of systems which do not exhibit the structure on which these laws and regularities are based.

Schrödinger jumps to that conclusion after reading that genes are specific molecules of which each cell generally contains no more than two copies. He had entered Vienna University in 1906, the year that Boltzmann died, and had been taught physics by Boltzmann's pupils. He remained deeply influenced by Boltzmann's thoughts throughout his life. According to Boltzmann's statistical thermodynamics, the behaviour of single molecules is unpredictable; only the behaviour of large numbers is predictable. In genetics, therefore, Schrödinger concludes, "we are faced with a mechanism entirely different from the probabilistic



N.W. Timoféeff-Ressovsky and (right) K.G. Zimmer: the co-authors with Delbrück of the 'Green Pamphlet', the paper which, in paraphrased form, constitutes the backbone of Schrödinger's book.

ones of physics". This difference forms the guiding theme of his book.

In the first chapter, Schrödinger illustrates the meaning of statistical thermodynamics by the examples of Curie's Law, of Brownian motion and diffusion, and of the $\sqrt{n}$ rule. His next two chapters, on hereditary mechanisms and on mutations, give brief popular introductions to textbook knowledge on these subjects available at the time. They reveal one vital misconception in Schrödinger's mind: "Chromosomes", he writes, "are both the law code and the executive power of the living cell". In fact, biochemists had shown that the executive power resides in enzyme catalysts, and in 1941 G.W. Beadle and E.L. Tatum discovered that single genes determine single enzymatic activities⁶; that discovery led to the one-gene-one-enzyme hypothesis, a most fruitful idea that had already been foreshadowed by the Cambridge biochemist and geneticist J.B.S. Haldane⁷. Schrödinger does not appear to have heard of this.

The next two chapters form the backbone of his book; they are called "The Quantum-Mechanical Evidence" and "Delbrück's Model Discussed and Tested", and are largely paraphrased versions of the paper by Timoféeff, Zimmer and Delbrück⁸. That paper covers 55 pages and is divided into four sections. The first section, by Timoféeff, describes the mutagenic effects of X-rays and γ -rays on the fruitfly *Drosophila melanogaster*. He shows that the spontaneous mutation rate of the fly is low, and that it is raised about five fold by a rise in temperature of 10°C. Ionizing radiations increase that rate as a linear function of the dose, independent of its time distribution, of the wavelength and of the temperature during irradiation.

The second section of the paper is by Zimmer and applies the target theory to Timoféeff's results. The number of mutations $x = a(1 - e^{-kD})$ where a and k are constants and D is the dose. Zimmer next asks whether the mutations had arisen by the direct absorption of quanta, by the passage of secondary electrons through a sen-

sitive volume or by the generation of ion pairs. If the dose is measured in roentgens, the number of quanta required to produce a given dose diminishes with diminishing wavelength. Thus direct absorption of quanta is inconsistent with the linear dependence of the mutation rate on the dose. The same applies to secondary electrons. Only the number of ion pairs is proportional to the dose, obviously, because that is how the dose is measured.

"Seymour Benzer, Maurice Wilkins and Gunther Stent have said that the book was decisive in drawing them into biology."

Zimmer therefore concludes that a single hit suffices for the production of one mutation and that this hit results either in the formation of an ion pair or a transition to higher energy.

The third section of the paper is by Max Delbrück and bears the title "Atomphysikalisches Modell der Mutation" ("A Model of Genetic Mutation Based on Atomic Physics"). Delbrück reminds us that the concept of the gene began as an abstract one, independent of physics and chemistry, until it was linked to chromosomes and later to parts of chromosomes which were estimated to be of molecular size. Since he and his colleagues had no means of discovering the chemical nature of genes directly, they attacked the problem indirectly by studying the nature and the limits of their stability and by asking if these were consistent with the knowledge that atomic theory has provided about the behaviour of well-defined assemblies of atoms.

Such assemblies can undergo discrete and spontaneous transitions of vibrational and electronic states. Vibrational transitions are very frequent and involve no chemical changes. From electronic transitions the assemblies may either revert to the ground state or reach a new equilibrium state after undergoing an atomic rearrangement, for example to a tautomeric form. The five-fold rise in spontaneous mutation frequency for a 10°C rise in

temperature leads Delbrück to derive an activation energy of $\sim 1.5\text{eV}$ and an average lifetime of a few years when half the molecules composing the gene will have undergone an electronic transition.

Delbrück then describes how, on average, X-rays lose energy to secondary electrons in portions of 30eV per ionization, which is $1,000 \times kT$ and 20 times the energy of activation of 1.5eV needed for a spontaneous mutation. However, to produce as much as 1.5eV, the ionization must not occur too far away from its target. We knew too little about the ways in which the energy of photoelectrons is dissipated to determine the absolute value of the dose needed to induce a mutation with a probability of unity, but that dose, expressed as the number of ionizations per unit volume, was likely to be about 10 – 100 times smaller than the number of atoms of the gene per unit volume. Delbrück now calculates that dose as follows.

A frequently observed X-ray mutation (eosin) occurs with a dose of 6,000r once in 7,000 gametes. Hence a probability of unity of its occurrence needs a dose of $42 \times 10^6\text{r}$. 1r produces $\sim 2 \times 10^{12}$ ion pairs in 1ml H_2O , whence $42 \times 10^6\text{r}$ produce $\sim 10^{20}$ ion pairs. Since 1ml H_2O contains $\sim 10^{23}$ atoms, this means that at least one in a thousand atoms becomes ionized. However, Delbrück cautiously refrains from concluding that a gene is likely to consist of a thousand atoms.

Schrödinger used Delbrück's result to point out "that there is a fair chance of producing a mutation when an ionization occurs not more than about 10 atoms away from a particular spot on the chromosome", but research published while Schrödinger was writing his book showed such calculations to be meaningless. In a paper that appeared in June 1944, Joseph Weiss pointed out that the biological effects of ionizing radiation are due principally to the generation of hydroxyl radicals and hydrogen atoms in the surrounding water⁹. Collinson, Dainton, Smith and Tazuke¹⁰, and independently Czapski and Schwartz¹¹, later discovered that the supposed hydrogen atoms were in fact

hydrated electrons¹². Hydroxyl ions and hydrated electrons have half lives of ~1 ms (assuming a concentration of 1 μ M H_2O_2) and 0.5 ms respectively, in which times they can diffuse to their targets even if they are generated more than a thousand atomic diameters away from them.

Delbrück concludes that it is premature to make the description of the gene any more concrete than the following:

We leave open the question whether the single gene is a polymeric entity that arises by the repetition of identical atomic structures or whether such periodicity is absent; and whether individual genes are separate atomic assemblies or largely autonomous parts of a large structure. i.e. whether a chromosome contains a row of separate genes like a string of pearls, or a physico-chemical continuum.

I found the Timoféeff-Zimmer-Delbrück paper and especially Delbrück's part, most impressive. Delbrück was a theoretical physicist whose interest in biology had been aroused by Niels Bohr's lecture "Light and Life", delivered in Copenhagen in 1932. In that lecture, Bohr had said:

The existence of life must be considered as an elementary fact that cannot be explained, but must be taken as a starting point in biology, in a similar way as the quantum of action, which appears as an irrational element from the point of view of classical mechanical physics, taken together with the existence of elementary particles, forms the foundation of atomic physics. The asserted impossibility of a physical or chemical explanation of the function peculiar to life would be . . . analogous to the insufficiency of the mechanical analysis for the understanding of the stability of atoms¹³.

The search for Bohr's elementary fact of life had fired Delbrück's imagination. He was only 29 years old, working as assistant to Otto Hahn and Lise Meitner in the Kaiser Wilhelm Institut für Chemie in Berlin and doing his biological work as a side line, but his paper shows the maturity, judgement and breadth of knowledge of someone who had been in the field for years. It is imaginative and sober, and its carefully worded predictions have stood the test of time. The paper won him a Rockefeller Fellowship to Pasadena to work with the *Drosophila* geneticist T.H. Morgan. There he met Linus Pauling with whom he published an important paper in 1940. That paper was an attack on the German theoretician P. Jordan who had advanced the idea that there exists a quantum-mechanical stabilizing interaction, operating preferentially between identical or near-identical molecules, which is important in biological processes such as the reproduction of genes. Pauling and Delbrück pointed out that interactions between molecules were now rather well understood and gave stability to two molecules of *complementary* structure in juxtaposition, rather than to two molecules with necessarily *identical* structures. Complementariness should be given primary

consideration in the discussion of the specific attraction between molecules and their enzymatic synthesis¹⁴. In 1937 the Cambridge geneticist and biochemist J.B.S. Haldane had made a similar suggestion: "We could conceive of a [copying] process [of the gene] analogous to the copying of a gramophone record by the intermediation of a negative, perhaps related to the original as an antibody is to an antigen"¹⁵. Schrödinger mentions neither of these important ideas.

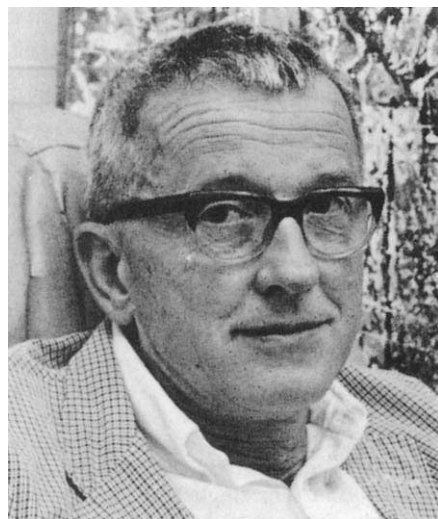
Schrödinger's last two chapters do contain his own thoughts on the nature of life. In "Order, Disorder and Entropy" he argues that "the living organism seems to be a macroscopic system which in part of its behaviour approaches to that purely mechanical (as contrasted with thermodynamical) conduct to which all systems tend, as the temperature approaches the absolute zero and the molecular disorder is removed". He comes to this strange conclusion on the ground that living systems do not come to thermodynamic equilibrium, defined as the state of maximum entropy. They avoid doing so, according to Schrödinger, by feeding on negative entropy. I suspect that Schrödinger got that idea from a lecture by Boltzmann on the Second Law, delivered before the Imperial Austrian Academy of Sciences in 1886:

Hence the general battle for existence of living organisms is not one for the basic substances — these substances are abundant in the air, in water and on the ground — also not for energy that every body contains abundantly, though unfortunately in a non-available form, but for entropy which becomes available by the transition of energy from the hot sun to the cold earth¹⁶.

Franz (later Sir Francis) Simon, then at Oxford, pointed out to Schrödinger that we do not live on $-T\Delta S$ alone, but on free energy¹. Schrödinger deals with that objection in the second edition of his book; he writes that he had realized the importance of free energy, but had regarded it as too difficult a term for his lay audience; to me this seems a strange argument, because the meaning of entropy is surely harder to grasp. Schrödinger's postscript did not satisfy Simon who pointed out to him in a letter that:

The reactions in the living body are only partly reversible and consequently heat is developed of which we have to get rid to the surroundings. With this irreversibly produced heat also flow small amounts (either + or -) of reversibly produced heat ($T\Delta S$), but they are quite insignificant and therefore cannot have the important effects on life processes which you assign to them¹.

In fact, it was known when Schrödinger wrote his book that the primary currency of chemical energy in the cells is ATP, and that the free energy stored in ATP is predominantly enthalpic. However, Schrödinger did not remove this misleading chapter from later editions.



Max Delbrück: predictions have stood the test of time.

The final chapter, "Is Life Based on the Laws of Physics?" reiterates and amplifies the central argument stated at the beginning of the book. According to Delbrück, Schrödinger writes, the gene is a molecule, but the bond energies in molecules are of the same order as the energy between atoms in solids, for example in crystals, where the same pattern is repeated periodically in three dimensions, and where there exists a continuity of chemical bonds extending over large distances. This leads him to the famous hypothesis that the gene is a linear one-dimensional crystal, but lacking a periodic repeat: an aperiodic crystal. A single such crystal, or a pair of them, direct the orderly process of life. Yet, according to Boltzmann's laws, their behaviour must be unpredictably erratic. Schrödinger concludes that:

We are faced with a mechanism entirely different from the probabilistic one of physics, one that cannot be reduced to the ordinary laws of physics, not on the ground that there is any 'new force' directing the behaviour of single atoms within an organism, but because the construction is different from any yet tested in the physical laboratory.

I wonder why Schrödinger did not adhere to Delbrück's much better formulation of "a polymeric entity that arises by the repetition of identical atomic structures". One could argue over the distinction between aperiodic and identical, but Delbrück could not have meant structures that are *completely* identical, since these could contain no information. Schrödinger does suggest that the genetic information might take the form of a linear code, analogous to the Morse code.

He argues that the nature of the gene allows only one general conclusion:

Living matter, while not eluding the laws of physics as established to date, is likely to involve other laws of physics hitherto unknown which, however, once they have been revealed, will form as integral a part of this science as the former.

Schrödinger is thus drawn to the same conclusion as Niels Bohr had been, apparently unknown to Schrödinger, twelve years earlier, and one that young physicists found equally inspiring.

Schrödinger next refers to a paper by Max Planck, "Dynamical and Statistical Laws". Dynamical laws control large-scale events such as the motions of the planets or of clocks. Clockworks function dynamically, because they are made of solids kept in shape by London-Heitler forces strong enough to elude disorderly heat motions at ordinary temperatures. An organism is like a clockwork in that it also hinges upon a solid: the aperiodic crystal forming the hereditary substance, largely withdrawn from the disorder of heat motion. The single cog of this clockwork is not of coarse human make, but is the finest piece ever achieved along the lines of the Lord's quantum mechanics. C.D. Darlington at Oxford had advised him that genes are likely to be protein molecules, as was then generally believed; Schrödinger quotes that, but does not mention that proteins are long chain polymers made up of some 20 different links that might form the kind of aperiodic patterns or linear code he had in mind. He must also have been unaware that the true chemical nature of that "finest piece" was actually published while he was writing his book. In January 1944 there appeared a paper by O.T. Avery, C.M. McLeod and M. McCarty which reported conclusive evidence that genes are made not of protein, but of DNA¹⁶. In the fullness of time, that discovery has led most scientists to the recognition that life can be explained on the basis of the existing laws of physics.

When I was invited to review the influence of *What is Life?* I accepted with the intention of doing honour to Schrödinger's memory. To my disappointment, a close study of his book and of the related literature has shown me that what was true in his book was not original, and most of what was original was known not to be true even when it was written. A.F. Huxley told me that even the charming contrast between the probable lifetime of a radioactive atom and a sparrow was misplaced. The incidence of death among a population of song birds hatched in any one Spring was known to follow exactly the same law as radioactive decay; for example, for the first ten years of its life, the probable lifetime of a robin is about one year, regardless of age. Older birds are too rare for statistics¹⁷. There is no reason to believe the lifetime of sparrows to be any more predictable than that of other birds. By contrast, my reading has raised even further my already great respect and admiration for Delbrück's analytical powers and scientific rigour, and for the prophetic and imaginative concepts formulated by J.B.S. Haldane and Linus Pauling, often long in advance of the rel-

evant discoveries. In retrospect, the chief merit of *What is Life?* is its popularization of the Timoféeff, Zimmer and Delbrück paper that would otherwise have remained unknown outside the circles of geneticists and radiation biologists.

The apparent contradictions between life and the statistical laws of physics can be resolved by invoking a science largely ignored by Schrödinger. That science is

"The apparent contradictions between life and the statistical laws of physics can be resolved by invoking a science largely ignored by Schrödinger. That science is chemistry."

chemistry. When Schrödinger wrote: "The regular course of events, governed by the laws of physics, is never the consequence of one well-ordered configuration of atoms, not unless that configuration repeats itself many times", he failed to realize that this is exactly how chemical catalysts work. Given a source of free energy, a well-ordered configuration of atoms in a single molecule of an enzyme catalyst can direct the formation of an ordered stereospecific compound at a rate of 10^3 – 10^5 molecules a second, thus creating order from disorder at the ultimate expense of solar energy. Haldane pointed this out in 1945, in his review of Schrödinger's book¹⁸.

Chemists could also have told him that there is no problem in explaining the stability of the large molecules of living matter that so much exercised him, because their bond energies range from 3eV upwards which corresponds to a half-life for each bond of at least 10^{30} years at room temperature. The difficulty resides in explaining how their aperiodic patterns are accurately reproduced in each generation. There is no mention of this central problem in Schrödinger's book; but its importance was recognized soon after Watson and Crick proposed their mechanism of DNA replication.

That replication is catalysed by an enzyme or system of enzymes that attach themselves to the end of the DNA double helix, unwind it, hold each parent strand rigidly in the conformation needed to catalyse the formation of a new chain link, move forward one step, catalyse the formation of the next link, and so on. The enzymes are large enough to arrest the random motions of the DNA chain, thus allowing an orderly process to take place in a single molecule without violating the known laws of physics^{19,20}. A system of enzymatic proof-reading and editing ensures that the error rate in DNA replication is between 10^{-9} and 10^{-10} , four to five orders of magnitude less than it would otherwise be.

Twenty five years after the Timoféeff,

Zimmer and Delbrück paper was published, H. Traut, then a graduate student working in Zimmer's laboratory, re-examined the evidence on which it was based and found the linear dose-response curve of Timoféeff to have been an artefact. He showed that the mutation rate of *Drosophila* germ cells varies widely at different stages of their development. If males are irradiated and then mated, the frequency of mutation among the offspring varies with the time that has elapsed between the two events, because the sperm that fertilizes a female five days after irradiation is at an earlier stage of development when it is irradiated than the sperm that fertilizes a female one day after irradiation. At all stages the dose-response curves are non-linear. Traut demonstrated that a linear response curve, similar to those observed by Timoféeff, is obtained by summing the different dose-response curves produced by matings during the first four days of irradiation²¹.

Zimmer comments:

This result removes one of the foundation-stones of the Green Pamphlet [as the Timoféeff, Zimmer & Delbrück paper became known]. Strangely enough, that does not seem to matter any more, for two reasons; (i) the concept of the gene and modern trends in genetic research as well as in radiation biology have changed considerably during thirty years, and (ii) the Green Pamphlet has served a useful purpose by helping to initiate these modern trends²². □

1. Yoxen, E.J. *Hist. Sci.* **17**, 17–52 (1979).
2. Timoféeff-Ressovsky, N.W., Zimmer, K.G. & Delbrück, M. *Nachrichten aus der Biologie der Gesellschaft der Wissenschaften Göttingen* **1**, 189–245 (1935).
3. Jacob, F. in *La Logique du Vivant*, 279 (Editions Gallimard, Paris, 1970).
4. Crick, F.H.C. *Br. Med. Bull.* **21**, 183–186 (1965).
5. Crick, F.H.C. quoted by R.C. Olby, *J. hist. Biol.* **4**, 119–148 (1971).
6. Beadle, G.W. & Tatum, E.L. *Proc. natn. Acad. Sci. U.S.A.* **27**, 499–506 (1941).
7. Haldane, J.B.S. in *Perspectives of Biochemistry* (eds Needham, J. & Green, D.E.) 1–10 (Cambridge University Press, 1937).
8. Waddington, C.H. *Nature* **221**, 318–321 (1969).
9. Weiss, J. *Nature* **153**, 748–750 (1944).
10. Collinson, E., Dainton, F.S., Smith, D.R. & Tazuke, S. *Proc. Chem. Soc.* 140–144 (1962).
11. Czapski, G. & Schwartz, H.A. *J. phys. Chem.* **66**, 471–479 (1962).
12. Dainton, F.S. *Chem. Soc. Rev.* **4**, 323–362 (1975).
13. Bohr, N. *Nature* **131**, 458–460 (1933).
14. Pauling, L. & Delbrück, M. *Science* **92**, 77–79 (1940).
15. Boltzmann, L. *Der zweite Hauptsatz der mechanischen Wärmetheorie*. (Sitzungsber. Kaiserl. Akad. Wiss. Wien, 1886).
16. Avery, O.T., McLeod, C.M. & McCarty, M. *J. exp. Med.* **79**, 137–158 (1944).
17. Lack, D. in *The Life of the Robin*, 117 (Pelican Books, London, 1943).
18. Haldane, J.B.S. *Nature* **155**, 375–376 (1945).
19. Kornberg, A. *DNA Replication* (Freeman, San Francisco, 1980).
20. Kornberg, A. *Supplement to DNA Replication* (Freeman, San Francisco, 1982).
21. Traut, H. in *Repair from Genetic Radiation Damage* (ed. Sobels, F.H.) 359 (Pergamon, London, 1963).
22. Zimmer, K.G. in *Phage and the Origins of Molecular Biology* (eds Cairns, J., Stent, G.S. & Watson, J.D.) 33–42 (Cold Spring Harbor Laboratory, New York, 1966).

M.F. Perutz is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. This article is adapted from his chapter in Schrödinger: Centenary Celebration of a Polymath, published this month by Cambridge University Press.

Minimal chaos and stochastic webs

A. A. Chernikov, R. Z. Sagdeev, D. A. Usikov, M. Yu Zakharov & G. M. Zaslavsky

Space Research Institute, Academy of Sciences of the USSR, Profsoyuznaya 84/32, Moscow 117810, USSR

Particles within a stochastic—or chaotic—web in phase space can be accelerated to high energies even by weak magnetic fields, and the properties of the web show similarities to the quasicrystal state. Examples of webs with five- and sevenfold symmetries are given. The establishment of stochastic webs and structural coverings of the phase plane may yield new and unexpected results.

STOCHASTICITY, or chaos, in a system is understood as the occurrence of statistical dynamics in the absence of random forces. This phenomenon is typical of most nonlinear systems, and progress in understanding chaos is one of the characteristic features of the modern physics of nonlinear phenomena. These developments have shed new light on such universal physical concepts as adiabatic invariants, integrals of motion and their stability, and the onset of turbulence. The principal non-integrability of the systems and stochastic diffusion under chaotic conditions result in numerous physical consequences, including the existence of instability regions in problems of celestial mechanics, the acceleration of particles, the violation of periodicity in macroscopic processes and the disruption of atomic quantum numbers in strong magnetic fields. Such phenomena have a considerable impact on our understanding of the physical universe. Even weakly chaotic behaviour may be accessible to observations if it spans astronomically large time-periods. Therefore, the problem of minimal conditions and regions of stochasticity is a significant one. Clearly the topic of chaos has wide scope and so it is useful to restrict our discussion to a particular physical problem.

In 1949 Fermi formulated the idea of stochastic acceleration of charged particles due to their collisions with randomly moving magnetic clouds¹. Actually, huge clouds, carriers of magnetic fields, have only small-scale turbulent field structure, while their large-scale distribution can be regular, reflecting their origin. This is a particular example in which chaotic particle dynamics even in regular (periodic) fields should play a principal role. Fermi understood this possibility clearly in formulating the Fermi–Pasta–Ulam problem². These authors attempted to illustrate numerically thermalization in a system of coupled nonlinear oscillators. Are the routes to chaos as numerous as are different physical systems? If one is restricted to hamiltonian systems only, the choice is rather poor and is determined by the fact that in the hamiltonian case only a very special kind of motion can change abruptly depending on the chosen initial conditions. For example, at sufficiently high energy the oscillations of a pendulum may transform into rotations. The curve that separates, in phase space, two types of motion but does not belong to either of them, is known as the separatrix. In the phase plane the separatrix has the form of one or several loops (cells) with one or more self-crossing points which are called saddles. Near these saddles the motion slows down and becomes very sensitive even to very weak perturbations. This is why a stochastic layer (a region of stochastic dynamics) arises at the unperturbed separatrix under the influence of arbitrarily small and even regular perturbations³. The thickness of the layer depends on the strength and character of the perturbation⁴. In the simplest case, the equation of motion which induces formation of the stochastic layer, is the equation of motion of the nonlinear oscillator disturbed by a plane wave:

$$\ddot{x} + \omega_0^2 \sin x = (\varepsilon \omega_0^2 / k) \sin(kx - \Omega t) \quad (1)$$

where $\varepsilon \ll 1$ is the dimensionless disturbance amplitude. The separatrix in phase space is destroyed whatever the magnitude

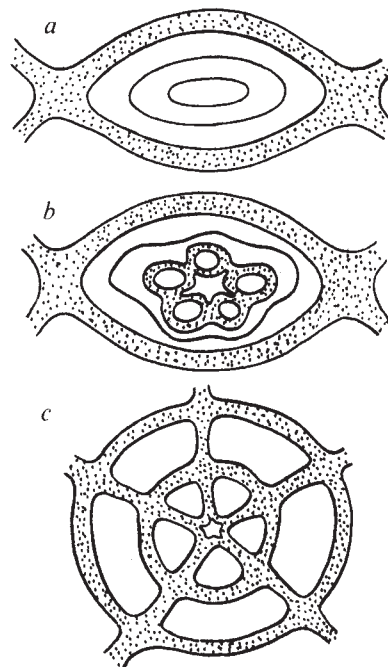


Fig. 1 Phase plane $(x, \dot{x})$ with thin stochastic regions (filled by dots) shown for various systems described here. The x -axis is horizontal, the $\dot{x}$ -axis is vertical and exact positioning of the axes is not essential. *a*, The stochastic layer formed in place of disrupted separatrix in equation (1) for non-resonant disturbance. Particle dynamics is stochastic inside the layer. Outside the layer but inside the separatrix cell reside the invariant curves, where Kolmogorov–Arnold–Moser theory is applicable. *b*, Emergence of substructure inside the separatrix cell under resonance conditions given by equation (2). Invariant curves are replaced by a new separatrix which possesses rotational symmetry and represents a thin filament with inner stochastic dynamics. Applicability of the KAM theory is severely restricted. *c*, Separatrix net for equation (3) under resonance condition (2) is unbounded on the phase plane. This net forms a stochastic web, along which diffusional wandering of the particles occurs.

of ε , and a region of stochastic dynamics is formed in its place (Fig. 1a). Invariant trajectories, which result from undisturbed trajectories through a weak deformation occur on both sides of the stochastic layer. A region of invariant curves covers practically the whole of the separatrix cell. This region is the very region of applicability of the Kolmogorov–Arnold–Moser (KAM) theorem⁵, that is, the theory of preservation of stable (in a certain sense) trajectories under the influence of the perturbation.

The example given by equation (1) is quite adequate for the general physical description of the onset of chaos and therefore it is very widely discussed in connection with various problems, for example, the disruption of magnetic surfaces in closed mag-

netic traps, particle dynamics in free-electron lasers, structural phase transitions in one-dimensional chains, fluctuations in superconducting connections and the acceleration of charged particles in plasmas.

Up until now, various modifications of the model given by (1) have not produced any essentially new qualitative changes in the picture of Fig. 1a. Our results modify this established concept of the structure of stochastic layers. Moreover, the new concept of the onset of chaos described below highlights an unexpected connection between the problem of minimal chaos in low-dimensional systems and the problems of plane tilings and of quasicrystal structures. The discovery of 'shechtmanites', which possess fivefold symmetry has had a considerable impact on the physics of crystals. Possible structures of shechtmanite are being actively explored. The 'tiler mapping' proposed below allows us to construct two-dimensional structures with arbitrary rotational quasi-symmetry. This establishes a new non-trivial interrelation between the problem of stability of dynamic systems and structures of quasicrystals.

Particle-wave interaction in a magnetic field

A qualitative change in the form of the region of chaos onset of the type shown in Fig. 1a may occur when the KAM theorem is violated. For the problem given in equation (1), this occurs when the resonance condition

$$\Omega - n\omega_0 = 0 \quad (2)$$

holds, where n is an integer (Fig. 1b). For certain values of ε : $1 > \varepsilon > \varepsilon_c$ a substructure arises within the separatrix cell (A.A.C. *et al.*, preprint, Space Research Institute, Moscow, 1986). It is also a region of stochastic dynamics that takes place in the disrupted inner separatrix. Therefore, for $\varepsilon > \varepsilon_c$ invariant curves are replaced by a new separatrix, which has rotational symmetry of order n ; we call this a web. It is 'disrupted' from the very beginning, the web represents a thin filament within which the dynamics are stochastic. The region of applicability of the KAM theory becomes much smaller because invariant curves remain either deep inside the web or in a region between the web and the stochastic layer. This example demonstrates the possibility of stochastic particle acceleration along web filaments. This is, however, still a restricted possibility since the size of the web is restricted.

Equation (1) seems to become simpler for particle motion in an external uniform magnetic field B_0 and for the motion in presence of a plane wave that propagates orthogonally to the magnetic field (along the x -axis):

$$\ddot{x} + \omega_0^2 x = (\varepsilon \omega_0^2 / k) \sin(kx - \Omega t) \quad (3)$$

where $\omega_0 = eB_0/mc$ is the cyclotron frequency.

This situation is typical of both laboratory and astrophysical plasmas. Stochasticity has a number of amusing and peculiar properties. We are accustomed to the fact that various solutions which are 'good' in simple models, are spoiled when more complex models are adopted. The situation changes, however, in the region of stochastic dynamics. Further complications of the model lead to an increase in the phase volume occupied by the stochastic region. Hence, to discover the condition for the onset of stochasticity, it is sufficient to study the simplest situations. In this sense, the problem given by equation (3) seems to represent the simplest case of particle dynamics in complicated electric and magnetic fields.

In spite of the apparent simplicity of equation (3), as compared to equation (1), the stochastic properties of the former are much more pronounced. This is due to the linear nature of the left-hand side of equation (3), which results in much greater degeneracy under the resonance condition (2). The hamiltonian which generates equation (3) is given by

$$H = \frac{1}{2}(\dot{x}^2 + \omega_0^2 x^2) + (\varepsilon \omega_0^2 / k^2) \cos(kx - \Omega t) \quad (4)$$

Motion in the vicinity of the resonance given by (2) can be

described in a standard way through specific action-angle variables (I, φ):

$$x = (2nI/\omega_0)^{1/2} \sin[(\varphi + \Omega t)/n]$$

$$\dot{x} = (2nI\omega_0)^{1/2} \cos[(\varphi + \Omega t)/n]$$

With these variables, equation (4) can be easily transformed to

$$H = H_0 + H_1 \quad H_0 = -\delta\omega I + (\varepsilon \omega_0^2 / k^2) J_n(k\rho) \cos \varphi$$

$$H_1 = (\varepsilon \omega_0^2 / k^2) \sum_{m \neq n} J_m(k\rho) \cos\left(\frac{m}{n}\varphi + \frac{(m-n)}{n}\Omega t\right) \quad (5)$$

where $\delta\omega = \Omega - n\omega_0$ is a shift from resonance, $\rho = (2nI/\omega_0)^{1/2}$ is the Larmor radius, and J_m is the Bessel function. In equation (5) H_1 includes an explicit time dependence at frequency $\geq \Omega/n \approx \omega_0$. Meanwhile, the unperturbed part, described by hamiltonian H_0 corresponds to modulation oscillations at a frequency of the order $\varepsilon\omega_0$. Thus, a high-frequency disturbance H_1 creates an exponentially thin stochastic layer in the vicinity of the separatrix of the undisturbed system H_0 (A.A.C. *et al.*, preprint, Space Research Institute, Moscow, 1986). Let us determine the position of the stochastic layer in the phase space for exact resonance ($\delta\omega = 0$). As follows from the determination of H_0 the equation for a singular point is:

$$J'_n(k\rho) \cos \varphi = 0 \quad J_n(k\rho) \sin \varphi = 0$$

which gives the following equations that determine the separatrix net:

$$\rho = \rho_s \quad (s = 1, 2, \dots) \quad \varphi = \pi(m + \frac{1}{2}) \quad (m = 1, 2, \dots) \quad (6)$$

where ρ_s are roots of the Bessel function: $J_n(k\rho_s) = 0$.

The separatrix net and the polar angle $\theta = \varphi/n$ are shown, on the phase plane ($x, \dot{x}$), in Fig. 1c for $n = 5$. This net is infinite in the phase plane and this implies that there is a significant possibility of diffusion *ad infinitum*. In other words, along filaments of the stochastic web, particles can be accelerated up to arbitrarily large energies. This phenomenon has some important features which deserve special attention.

It is known that when the number of degrees of freedom $N \geq 3$, there exists a certain separatrix net that fills the phase space and is covered by a thin stochastic layer. This net forms a stochastic web, along which particles diffuse. The existence of such a web was predicted by Arnold⁸. It has a topological origin since, for $N \geq 3$, invariant tori do not divide the phase space and therefore invariant tori intersect. They form a self-crossing separatrix net. For this reason stochastic filaments running along separatrices cannot be divided by invariant curves. The example of equation (3) gives us a similar picture of universal diffusion along the web filaments, but here the number of degrees of freedom $N = 3/2$ (half degree of freedom is commonly attributed to a temporary periodic disturbance). Since the case $N = 1$ is always integrable, our result implies the existence of a diffusion of the Arnold type for the minimal dimension of a non-integrable system.

In a more general case, temporal variations of the disturbance in the hamiltonian (4) contain not only one mode of frequency Ω , but, say, a set of modes with frequencies $n_1\Omega$. Then the more general resonance condition

$$n_1\Omega - n\omega_0 = 0$$

replaces equation (2); the frequency ratio Ω/ω_0 is rational. For incommensurable frequencies ω_0 and Ω one can form a suitable rational approximation

$$\left| \frac{\omega}{\Omega} - \frac{n_1}{n} \right| < \varepsilon_1 < \varepsilon$$

which also generates a stochastic web (n_1 and n are integers). The greater n the greater is the central region containing invariant curves, where KAM theory is applicable. In other words, as the order n of the nearby resonance increases, the stochastic web expands tending to infinity.

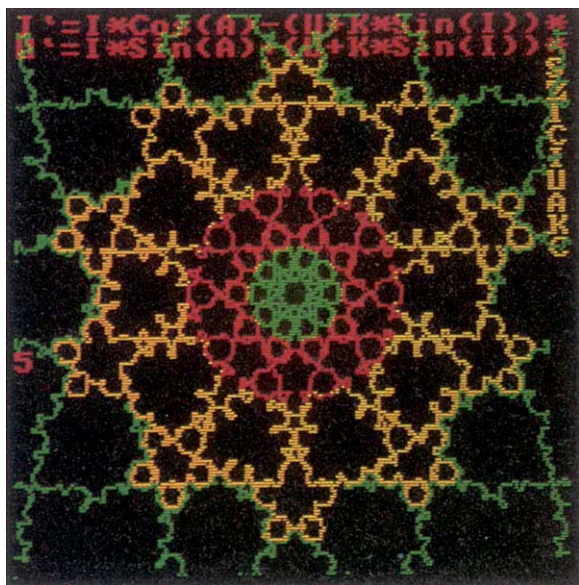


Fig. 2 Phase plane (u, v) and a stochastic web on it that is generated by the mapping. See equation (10) for $\alpha = 2\pi/5$ and $K = 0.7$. The web possesses a fifth order central symmetry. In the web cells particle dynamics are non-stochastic. The stochastic web presents an example of fractal structure. All points of the web belong to a single particle trajectory. Different time intervals of the trajectory are shown by different colours.

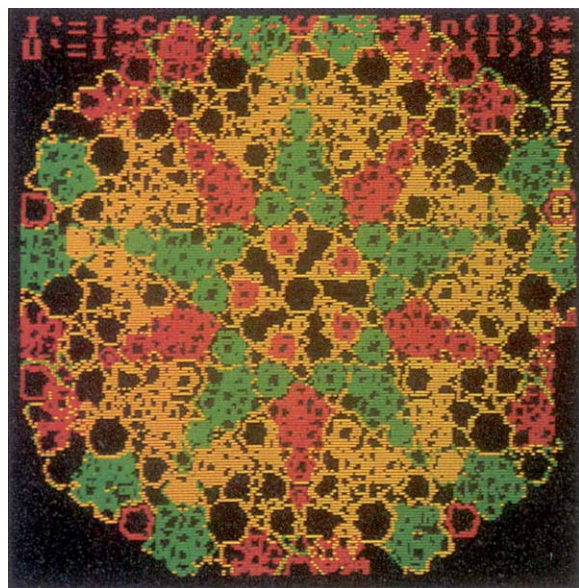


Fig. 3 Disruption of the short-range order in the stochastic web with growth of the nonlinear parameter K . Thickness of the stochastic layer becomes comparable to the web cell size at $K = 0.9$. Particle diffusion along the web filaments becomes more effective and the shape of the web cells changes quasi-randomly.

The following is a direct consequence of this result. All adiabatic invariants of the problem of the type given by (3) may be only conditional. This means that there exist finite (though small) regions in the phase space such that adiabatic invariants may vary to an arbitrary degree for the states belonging to these regions. These regions are the filaments of the stochastic web.

A uniform web

A more exact analysis⁴ can yield the thickness of the stochastic web. It turns out that its thickness decreases exponentially with the distance from the centre, and therefore diffusion weakens

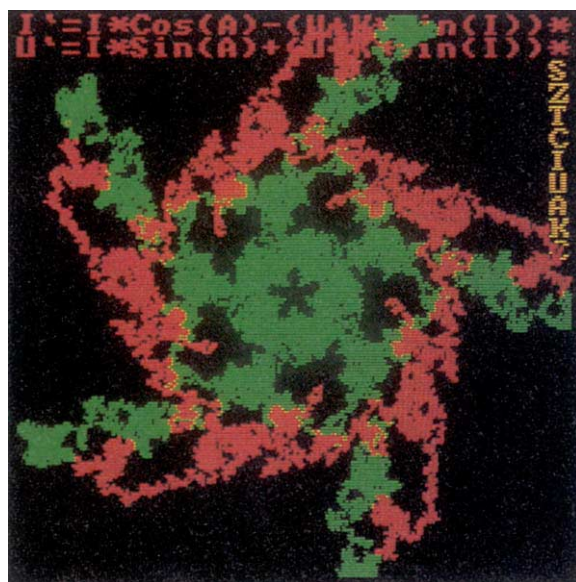


Fig. 4 A fractal tree formed for strong particle diffusion at $K = 20$. The diffusion can be described by the Fokker-Planck-Kolmogorov equation.

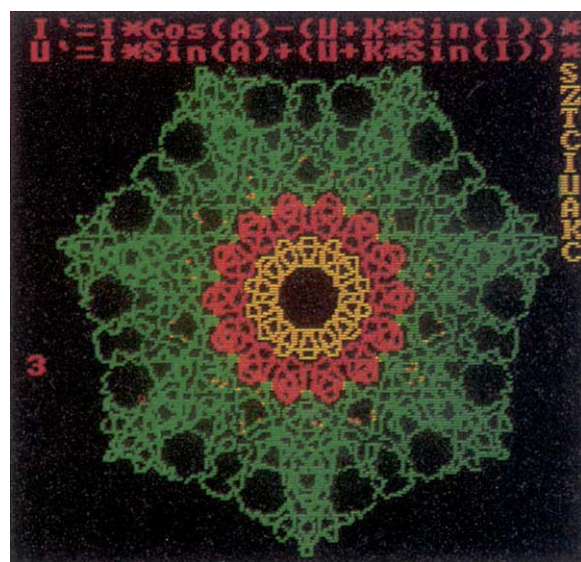


Fig. 5 An example of a stochastic web generated by the map (10) for $\alpha = 2\pi/7$ and $K = 0.5$. The pattern has seventh order central symmetry.

with the growth of the action. This fact somewhat weakens our result on universality of the diffusion mechanism for the minimal number of degrees of freedom, $N = \frac{3}{2}$. Can a uniformly 'thick' stochastic web cover the whole phase plane for $N \geq \frac{3}{2}$? The answer is affirmative and this directly implies the existence of a universal mechanism of stochastic particle acceleration, of the Fermi type, for an arbitrary number of degrees of freedom $N \geq \frac{3}{2}$. Now we turn to the example which proves this.

Equation (3) includes only one plane wave as a disturbance acting on a particle moving in magnetic field. More generally, we may consider a disturbing wave packet that contains a great number of plane waves

$$\ddot{x} + \omega_0^2 x = \frac{e}{m} \sum_k E_k \sin(kx - \omega_k t) \quad (7)$$

Particle dynamics become more complicated the more complex

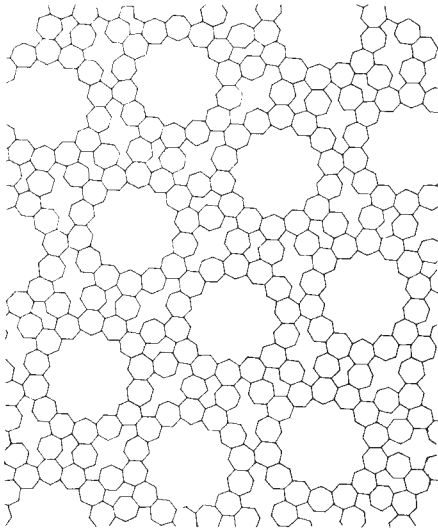


Fig. 6 An example of non-periodic covering that can be formed from points of the stochastic web having seventh order central symmetry. The covering consists of four elements: the regular heptagon, the fourteen-rayed star and two additional elements.

the structure of the wave packet. Therefore, we chose a certain special form of the wave packet:

$$E_k = \text{const} = E_0 \quad k = \text{const} = k_0 \quad \omega_k = n\Delta\omega (n = 0, \pm 1, \dots)$$

Here $\Delta\omega$ is the frequency separation of the neighbouring modes in the wave packet. The physical situation, in which such wave packets are realized, corresponds to particle motion at speeds below the packet group velocity^{7,8}. For a very large number of modes equation (7) reduces to

$$\ddot{x} + \omega_0^2 x = (K/k_0 T^2) \sin k_0 x \sum_{n=-\infty}^{+\infty} \delta(t/T - n) \quad (8)$$

with $T = 2\pi/\Delta\omega$ and

$$K = eE_0 k_0 T^2 / m \quad (9)$$

Equation (8) describes the behaviour of a linear oscillator under the action of a periodic sequence of kicks (δ -impulses). Their amplitude is determined by the dimensionless parameter K , which is equivalent to the disturbance parameter ε in equation (3). The kicks are separated by the interval T . This allows us to replace equation (8) by a finite-difference equation that relates amplitudes and phases between two consecutive collisions. It is convenient to introduce dimensionless variables

$$u = k\dot{x}/\omega_0; \quad v = -k_0 x$$

Let their values immediately before the n th kick be denoted by (u_n, v_n) . A recurrence relation easily follows from equation (8)

$$\hat{M}_\alpha: \begin{cases} u_{n+1} = [u_n + (K/\alpha) \sin v_n] \cos \alpha + v_n \sin \alpha \\ v_{n+1} = -[u_n + (K/\alpha) \sin v_n] \sin \alpha + v_n \cos \alpha \end{cases} \quad (10)$$

where $\alpha = \Delta\omega T$.

The mapping given in equation (10) (derived in ref. 10) has many remarkable properties. It describes a dynamic system with $N = \frac{3}{2}$ degrees of freedom, but it also contains all properties of stability and stochasticity that are relevant to equations (1) and (3). However, the most outstanding feature of the mapping $\hat{M}_\alpha$ is that it contains all known specific peculiarities of the hamiltonian dynamics of few-dimensional systems: all known types of bifurcations, strong and weak chaos, stochastic web and Arnold type diffusion. Some other points should be added: (i) equation (8) has the minimal possible dimension which prohibits integrability ($N = \frac{3}{2}$); (ii) equation (8) has regions of chaotic dynamics for disturbance of arbitrary magnitude K ;

and (iii) the stochastic web that arises under resonances

$$\alpha_q = \omega_0 T = 2\pi p/q \quad (11)$$

(p and q are integer) covers the whole phase plane and has a uniform thickness. The resonance condition (11) can be written as $q\omega_0 - p\Delta\omega = 0$ which is analogous to equation (2). This last property is especially important for numerous applications, since it implies an efficient diffusion in various regions of the phase space.

An example of a stochastic web for $\alpha_5 = 2\pi/5$ is given in Fig. 2. This web consists of the points of a single trajectory beginning within the web. Clearly, Fig. 2 depicts only a part of this web which not only covers the entire infinite phase plane, but also has a finer cell structure that we have not yet resolved due to finite computation time. A striking property of the stochastic web is the symmetry centre of the fifth order. This peculiarity deserves a special discussion.

Structural symmetry of the web

For any resonant value α_q that satisfies (11) the separatrix net has rotational quasi-symmetry of order q . The symmetry is more exact when the parameter K is smaller, that is when the web is thinner. General results on separatrix thickness^{4,9} imply the following estimate for the thickness of the stochastic web for weak disturbances:

$$\Delta \sim \exp(-\text{const}/K), \quad K \ll 1 \quad (12)$$

The smaller the disturbance parameter K , the more ordered are the web cells. Nevertheless, the web generally has fractal geometry. Its self-similarity properties develop with its size, that is with growth of the time of diffusion. The web walls are also fractal, as is any stochastic layer.

The two following diagrams, Figs 3 and 4, illustrate the variation of structures with the size of the disturbance, that is, the value of K . The stochastic web expands, covering a still greater part of the phase space. For large values of K , a coarse-grained region occupied by the particle trajectory represents a fractal tree having rotational symmetry of the fifth order (Fig. 4). Note that symmetry always means a quasi-symmetry, as the disturbance generates the symmetry on the one hand but, on the other hand, causes its slight deformation.

The mapping $\hat{M}_\alpha$ given by (10) is a generator of phase plane coverings with quasi-symmetry of any order $q = 2\pi/\alpha$ (q is an integer). An example of covering with a symmetry of the seventh order ($\alpha_7 = 2\pi/7$), is given in Fig. 5. Types of symmetry that arise for different α_q suggest the form of the equations of motion required to describe various structures that arise in nonlinear dynamic systems. Examples of such systems are: some astrophysical objects with spiral structures, numerous problems of turbulence (Benard cells) and quasi-crystal structures in solids^{10,11}. One of the outstanding problems with quasi-crystals concerns the structure of the atom lattice having fivefold symmetry. The mapping $\hat{M}_\alpha$ with $\alpha = 2\pi/q$ solves this question for any arbitrary order of symmetry. To do this, one should find the 'undressed' separatrices and place the atoms in the minima of the potential. This procedure has been described already¹². A universal characteristic of symmetry in nature manifests itself in a quite unexpected correspondence between the motion of a particle in a magnetic field plus a wave-packet field, and a non-trivial model of atomic ordering in quasi-crystals.

Conclusions

Many formal problems of the theory of dynamic systems turn out to have important practical applications. An example is the accuracy of conservation of adiabatic invariance in a magnetic field since this problem is central to initial attempts to estimate the confinement time for particles in magnetic traps. Similarly, the solution of the problem of the existence of the stochastic web allows one to establish the limits of adiabatic invariance and analyse the origin of ultra-high energy cosmic ray particles.

The phase space of the simplest dynamic systems (see equations (1), (3) and (10)) turns out to be remarkably rich in possible realizations of particle dynamics. Regular trajectories that are determined by dynamic invariants always exist. Also, however, there are always stochastic trajectories lacking the corresponding invariant. Stochastic dynamics can arise in exponentially small regions of phase space that are determined by the disturbance parameter given in equation (9). 'Minimal' stochasticity is determined by the minimal region (see equation (12)) and exists for arbitrary disturbances and any number of degrees of freedom >1 . 'Minimal' chaos takes the form of a stochastic web that arises in place of the separatrix net. The net can have various structural quasi-symmetries. Our results suggest the existence of a similar stochastic web in problems of hydrodynamic structures and structural modification of solids.

Received 25 September 1986; accepted 3 February 1987.

1. Fermi, E., *Phys. Rev.* **75**, 1169–1174 (1949).
2. Fermi, E., Pasta, J. R. & Ulam, S. *Studies of Nonlinear Problems* (Los-Alamos Rept LA-1940, 1955).
3. Filonenko, N. N., Sagdeev, R. Z. & Zaslavsky, G. M. *Nuclear Fusion* **7**, 253–266 (1967).
4. Zaslavsky, G. M. & Filonenko, N. N. *Zh. éksp. teor. Fiz.* **54**, 1590–1602 (1968).
5. Arnold, V. I. *Mathematical Methods of Classical Mechanics* (Nauka, Moscow, 1979).
6. Arnold, V. I. *Dokl. Akad. Nauk SSR* **156**, 9–12 (1964).
7. Zaslavsky, G. M. *Zh. éksp. teor. Fiz.* **88**, 1984–1995 (1985).

In particular, a direct correspondence can be established¹² between the web structure illustrated in Fig. 2 and the non-periodic tiling of Penrose³. An example of non-periodic tiling which consists of regular heptagons, fourteen-pointed stars and two additional elements is given in Fig. 6.

One more aspect of equations (3) and (8) is worth mentioning. These equations describe the interaction of two different types of symmetry of particle motion: rotational symmetry due to rotation in magnetic field and translational symmetry along x due to the field of the wave packet. These symmetries cannot be transformed one into another. As a result, for example, the paradox arises of the disappearance of Landau damping in the magnetic field¹⁴. The understanding of many related questions is associated with the nature of the non-trivial interaction of these two types of symmetry.

8. Zaslavsky, G. M. & Chernikov, A. A. *Zh. éksp. teor. Fiz.* **89**, 1632–1647 (1985).
9. Zaslavsky, G. M., Zakharov, M. Yu., Sagdeev, R. Z., Usikov, D. A. & Chernikov, A. A. *Zh. éksp. teor. Fiz.* **91**, 500–516 (1986).
10. Shechtman, D., Blech, I., Gratias, D. & Chan, J. W. *Phys. Rev. Lett.* **53**, 1951–1953 (1984).
11. Nelson, D. R. & Halperin, B. I. *Science* **229**, 233–238 (1985).
12. Zaslavsky, G. M., Zakharov, M. Yu., Sagdeev, R. Z., Usikov, D. A. & Chernikov, A. A. *Pisma Zh. éksp. teor. Fiz.* **44**, 349–353 (1986).
13. Penrose, R. *Bull. Inst. Math. Applic.* **10**, 266–271 (1974).
14. Sagdeev, R. Z. & Shapiro, V. D. *Pisma Zh. éksp. teor. Fiz.* **17**, 389–393 (1973).

Tests of the helix dipole model for stabilization of α -helices

Kevin R. Shoemaker*, Peter S. Kim*†, Eunice J. York‡, John M. Stewart‡ & Robert L. Baldwin*

* Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA

‡ Department of Biochemistry, University of Colorado School of Medicine, Denver, Colorado 80262, USA

Charged groups play a critical role in the stability of the helix formed by the isolated C-peptide (residues 1–13 of ribonuclease A) in aqueous solution. One charged-group effect may arise from interactions between charged residues at either end of the helix and the helix dipole. We report here that studies of C-peptide analogues support the helix dipole model, and provide further evidence for the importance of electrostatic interactions not included in the Zimm–Bragg model for α -helix formation.

THE C-peptide (residues 1–13) and S-peptide (residues 1–20) of RNase A show significant α -helix formation in aqueous solution near 0°C^{1–3}. Nevertheless, predictions of helix stability⁴ based on the Zimm–Bragg equation⁵ and using host–guest data⁶ indicate that helix formation should be negligible for all short peptides (≤ 13 residues), regardless of amino-acid sequence or temperature. Another surprise from studies of helix formation by C-peptide is the charged-group effect: the pH dependence of C-peptide helix content demonstrates the importance of ionizable groups².

Previous work⁴ has shown that the charged-group effect can be assigned to Glu 2[−] and His 12⁺. This result suggests that interactions with the helix dipole are important. Glu 2[−] and His 12⁺ are close to the positive and negative poles of the α -helix macrodipole, respectively, and can make favourable charge-dipole interactions that stabilize the helix.

The approach that we have taken is to study analogues of C-peptide that contain single amino-acid substitutions. Two preliminary generalizations can be made⁴: first, single amino-acid substitutions often have a detrimental effect on helix stability, although some which increase helix stability have also

been found; and second, the same amino-acid substitution can have opposite effects on helix stability depending on its location in the helix. With these points in mind, we have designed a C-peptide analogue which shows greater helix stability than C-peptide itself. Other analogues based on this reference peptide have been studied in order first to confirm the importance of Glu 2[−] and His 12⁺ and then to test the helix dipole model.

Reference peptide with increased helix stability

The sequences of native C-peptide and reference peptide III are given in Fig. 1 legend. Reference peptide III combines two amino-acid substitutions in the C-peptide sequence, Lys 1 $\rightarrow$ acetylAla and Glu 9 $\rightarrow$ Leu, known from earlier work⁴ to increase helix stability. The other two substitutions, Gln 11 $\rightarrow$ Ala and Met 13 $\rightarrow$ Ala, were made for ease of synthesis and purification⁴.

The far-ultraviolet (UV) CD (circular dichroism) spectrum of peptide III under optimal helix-forming conditions (pH 5.2, 3°C) is characteristic of partial helix formation (Fig. 1a). Minima occur at 222 nm (α -helix $n \rightarrow \pi^*$ transition) and 205 nm (mixture of 2 bands: α -helix $\pi \rightarrow \pi^*$ transition and random coil $\pi \rightarrow \pi^*$ transition)⁷. Recent work (ref. 8 and K.R.S. *et al.*, unpublished data) indicates that the change in $[\theta]_{222}$ (mean residue ellipticity at 222 nm) for complete helix formation by C-peptide at 3°C is near 30,000 deg cm² dmol^{−1}, when the baseline value

† Present address: Whitehead Institute, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA.

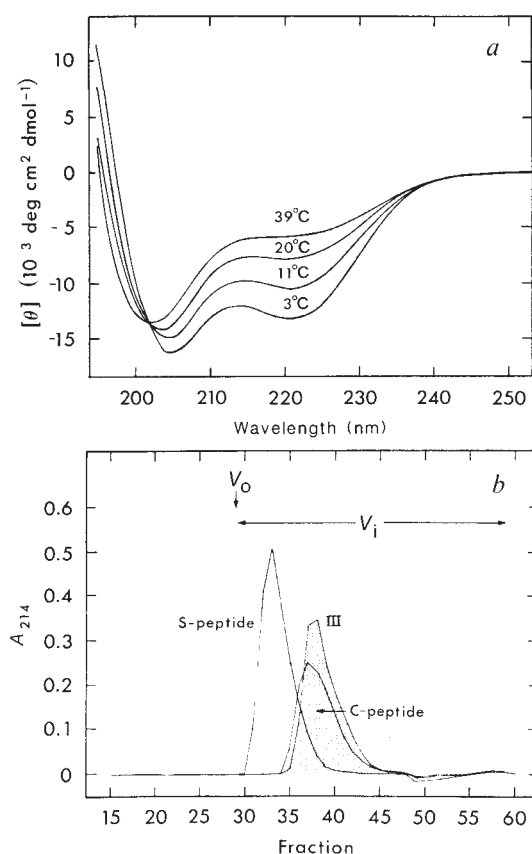


Fig. 1 *a*, Far-UV CD spectra of reference peptide III at various temperatures (pH 5.2, 0.1 M KF). The sequence of native C-peptide, obtained by cyanogen bromide cleavage at Met 13, is Lys-Glu-Thr-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Hse-lactone. The sequence of peptide III is acetyl-Ala-Glu-Thr-Ala-Ala-Lys-Phe-Leu-Arg-Ala-His-Ala-CONH₂. *b*, Gel filtration chromatography of peptide III (shaded; relative molecular mass M_r 1,398), C-peptide carboxylate (M_r 1,517), and S-peptide(1-19) (M_r 2,095) on Sephadex G-25 superfine. The conditions approximate those used for most CD measurements: 0.1 M NaCl, 1 mM sodium citrate, 1 mM sodium phosphate, 1 mM sodium borate, pH 5.2; 5 °C; peak fraction peptide concentration 20–30 μM . C-peptide has been shown by sedimentation equilibrium to be monomeric in this concentration range¹. Peptide elution was monitored by absorbance at 214 nm. V_0 and V_i mark the void volume and included volume, respectively.

Methods. Peptide synthesis and purification procedures have been described⁴. Both the Stewart Mark V and Biosearch 9500 automatic synthesizers were used. Peptide-resins were acetylated or succinylated with acetic anhydride or succinic anhydride, respectively, in dimethyl formamide containing an equivalent of triethylamine. CD methods were as described^{2,4} and CD spectra were recorded on two instruments: a Jasco J-500A spectropolarimeter in the laboratory of J. T. Yang (University of California Medical School, San Francisco) and an Aviv 60DS spectropolarimeter in the laboratory of P. S. Kim (Whitehead Institute, Cambridge, Massachusetts).

for 0% helix is estimated as +3,000 deg cm² dmol⁻¹. Thus, at pH 5.2 and 3 °C, peptide III shows ~50% helix content, substantially greater than either C-peptide² or reference peptide II⁴.

Figure 1*a* also indicates that the helix formed by peptide III unfolds with increasing temperature. This behaviour has been observed for C-peptide and all analogues studied to date^{2,4}. Two features of these CD spectra are consistent with a helix $\leftrightarrow$ random coil equilibrium: the presence of an isodichroic point at 202 nm and the shift in wavelength of the 205 nm minimum to 201 nm at higher temperatures.

Helix formation by peptide III is monomolecular, as judged by its mobility on a gel filtration column (Fig. 1*b*) and the lack

Table 1 Effect of NH₂-terminal charge on helix content in the absence of Arg 10

NH ₂ -terminal residue	$[\theta]_{222}$ (deg cm ² dmol ⁻¹ ; pH 5.2, 3 °C, 0.1 M NaCl)
succinylAla 1 (-1)	-17,400
acetylAla 1 (0)	-12,600
Ala 1 (+1)	-5,900
Lys 1 (+2)	-4,200

These peptides contain the substitution Arg 10 $\rightarrow$ Ala; otherwise, their sequences after residue 1 are the same as that of reference peptide III (see Fig. 1 legend). Charge of residue 1 in parentheses.

Table 2 Residue frequencies at ends of helices in proteins

NH ₂ -terminus			COOH-terminus		
Residue	<i>P</i>	Rank	Residue	<i>P</i>	Rank
Glu	2.44	1	Lys	1.83	1
Asp	2.02	2	His	1.77	2
His	0.73	11	Arg	1.25	5
Lys	0.66	16	Glu	1.24	6
Arg	0.44	20	Asp	0.61	16

The frequencies of helical boundary residues were obtained from 152 helices in 29 proteins. Adapted from Table VI of ref. 28.

P, normalized frequency of three helical end residues.

of any significant concentration dependence of $[\theta]_{222}$ in the range 10–40 μM (data not shown). The equilibrium between helix and random coil is mobile, and has not reached 100% helix in these studies. Therefore, if the helix were stabilized by oligomer formation, $[\theta]_{222}$ would vary with peptide concentration. Sedimentation equilibrium measurements have shown that C-peptide begins to form aggregates only in a much higher concentration range (>1 mM)¹.

Assignment of the charged-group effect

Figure 2*a* shows pH titrations of helix content for peptide III at 3 °C, where strong helix formation is observed, and at 45 °C, where the helix is largely melted out. In agreement with earlier findings⁴, ionization of either Glu 2 (pH 2 $\rightarrow$ 5) or His 12 (pH 8 $\rightarrow$ 5) results in a sharp increase in helix content. In this peptide, Glu 2 and His 12 are the only groups which titrate in these respective pH ranges. Residue replacement has been used to confirm these assignments: see the notation explained in Fig. 2 legend for peptide III analogues. Peptides III (Glu 2 $\rightarrow$ Ala) and III (His 12 $\rightarrow$ Ala) each show the expected drop in helix content at pH 5.2 and loss of either the acidic or basic branch of the pH titration curve (Fig. 2*b* and *c*)⁹.

To a first approximation, the helix-stabilizing effects of Glu 2⁻ and His 12⁺ are additive, suggesting that they are exerted independently of one another. Approximately the same change in helix content occurs on titrating His 12 in III (Glu 2 $\rightarrow$ Ala) as in peptide III with Glu 2 present, and the same is true of titrating Glu 2 in III (His 12 $\rightarrow$ Ala) as compared to peptide III.

NH₂-terminal charge and helix stability

The helix dipole model explains the charged-group effect by postulating that Glu 2⁻ and His 12⁺ interact with the α -helix dipole and thereby stabilize the helix by position-dependent electrostatic interactions. To test this model, we have varied systematically the charge on the NH₂-terminal residue of reference peptide III, from +2 in Lys 1 (first residue in native C-peptide) to -1 in succinylAla 1. Because solid-phase synthesis¹⁰ proceeds from the COOH-terminus to the NH₂-terminus, it was convenient to split the resin after Glu 2 coupling for addition of different NH₂-terminal residues.

Figure 3*a* shows pH titrations of helix content for these

Fig. 2 The pH dependence of mean residue ellipticity (222 nm) for *a*, reference peptide III, *b*, III (Glu 2→Ala), and *c*, III (His 12→Ala). The measurements were made at 3 °C (●) and 45 °C (○), 0.1 M NaCl. Analogues based on reference peptide III are denoted by giving the substitution: for example, III (Glu 2→Ala) is reference peptide III with alanine in place of Glu 2.

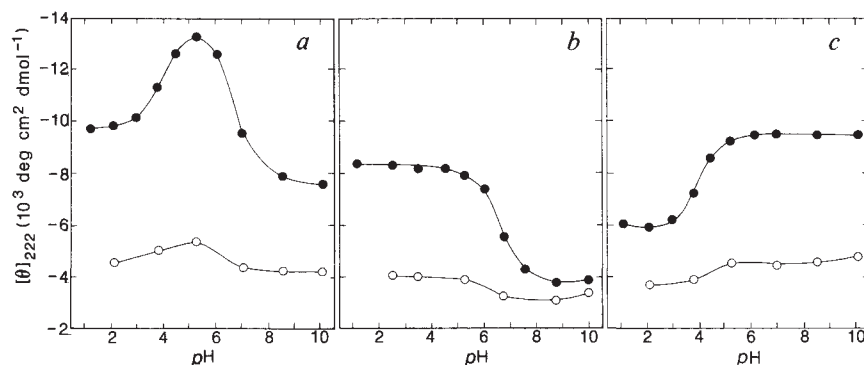
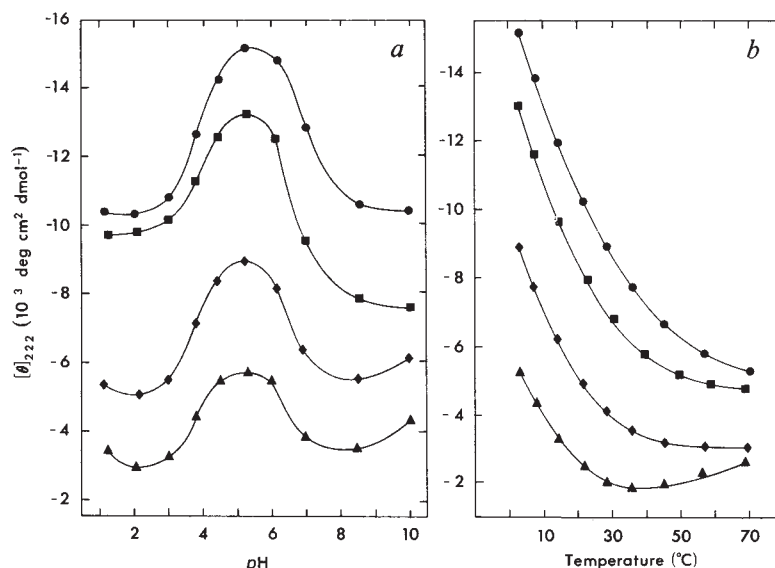


Fig. 3 *a*, The pH dependence of mean residue ellipticity (222 nm) at 3 °C, 0.1 M NaCl for reference peptide III (■), III (acetylAla 1→succinyl Ala) (●), III (acetylAla 1→Ala) (◆), and III (acetylAla 1→Lys) (▲). *b*, Temperature dependence of mean residue ellipticity (222 nm) for the same peptides at pH 5.2, 0.1 M NaCl. Peptides are denoted by the same symbols as in *a*.



peptides at 3 °C. All four peptides have Glu 2 and His 12, and they differ only in the charged residue at the NH₂-terminus. At pH 5.2, where the helix content is maximal and both Glu 2 and His 12 exist in their ionized forms, there is a progressive change in helix content with charge on the NH₂-terminus: succinylAla 1 (−1) > acetylAla 1 (0) > Ala 1 (+1) > Lys 1 (+2). This order of helix content agrees with predictions of the dipole model. As Fig. 3*b* indicates, helix content at 3 °C is correlated with thermal stability of the helix at pH 5.2.

Helix content has also been measured as a function of salt concentration for this set of peptides. Figure 4 shows that helix stability is affected by an electrostatic interaction involving the NH₂-terminus, as the nature of the interaction depends on the NH₂-terminal charge and the interaction is screened by salt. Differences in helix content are largest at low salt concentration. Increasing salt concentration favours helix formation when residue 1 is positively charged (Lys 1), but decreases helix content when residue 1 is negatively charged (succinylAla 1). As interpreted by the helix dipole model, NaCl is screening charge dipole interactions. As expected, the other two peptides in this series show behaviour intermediate between that of the Lys 1 and succinylAla 1 peptides (Fig. 4). Studies with other salts in the Hofmeister series indicate that the effects in this range of salt concentration result chiefly from generalized ion shielding and not specific ion interactions (data not shown).

Studies of peptides with the substitution Arg 10→Ala (manuscript in preparation) indicate that a Glu 2[−] ⋯ Arg 10⁺ ion pair, shown clearly in the refined X-ray and neutron crystal structure of intact RNase A¹¹, may also be present in the helix formed by the isolated C-peptide. To determine whether this ion pair is somehow involved in the effect demonstrated here, in which the NH₂-terminal charge strongly affects helix stability,

we have studied a similar set of peptides which contain the additional substitution Arg 10→Ala. Helix contents at pH 5.2, 3 °C, are summarized in Table 1. Both these data and thermal melting profiles (data not shown) indicate that the effect of NH₂-terminal charge is nearly the same in the presence or absence of the postulated Glu 2[−] ⋯ Arg 10⁺ ion pair.

Direct ionization at the NH₂ terminus

According to the helix dipole model, direct ionization of charged groups at the NH₂-terminus should affect helix stability. The peptide III (acetylAla 1→Ala, His 12→Ala) contains a free α-NH₂ group. As His 12 also ionizes above pH 5 (*pK_a* ~ 6.7⁴) and thereby alters helix content, His 12 has been replaced by alanine. Figure 5*a* shows pH titration of helix content at 3 °C, which demonstrates the effects of ionization of Glu 2 and the α-NH₂ group. As predicted, helix content increases as the protonated α-NH₂ group loses its positive charge (pH > 6.5).

Similarly, direct ionization of an NH₂-terminal succinyl group was studied in III (acetylAla 1→succinylAla, Glu 2→Ala). Figure 5*b* shows that helix content decreases with protonation of the negatively charged succinyl group. Here, alanine replaces Glu 2, which ionizes with *pK_a* ~ 3.8⁴.

Comparison with other work

Our results strongly support the previous suggestion⁴ that interactions between charged groups and the α-helix macrodipole influence the stability of the helix formed by C-peptide and its analogues. An early indication that these effects are energetically important was provided by the study of Chou and Fasman¹² on the distribution of charged amino acids near the termini of α-helices in proteins. As shown in Table 2, acidic residues occur preferentially near the NH₂-terminus but not near the COOH-

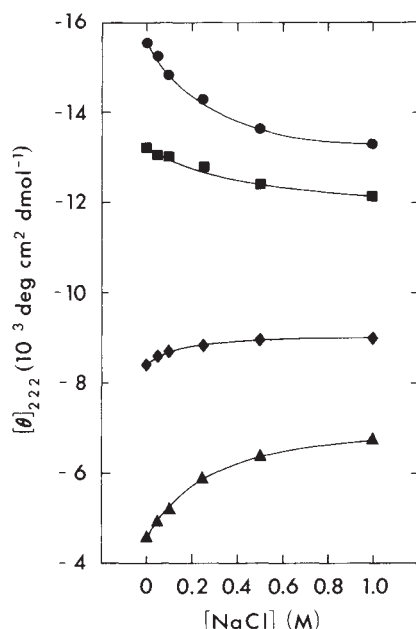


Fig. 4 Salt dependence of mean residue ellipticity (222 nm) at pH 5.2, 3 °C for the peptides described in Fig. 3 and denoted by the same symbols. Throughout this range of salt concentration, helix formation remains temperature dependent and occurs in a monomolecular reaction (data not shown).

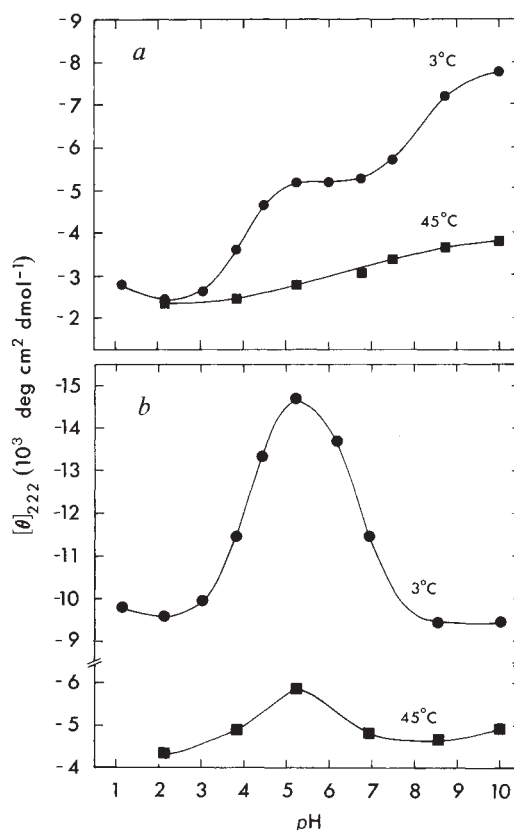


Fig. 5 The pH dependence of mean residue ellipticity (222 nm) for *a*, III (acetylAla 1 → Ala, His 12 → Ala) and *b*, III (acetylAla 1 → succinylAla, Glu 2 → Ala). The measurements were made at 3 °C (●) and 45 °C (■), 0.1 M NaCl.

terminus of α -helices, whereas the reverse is true of basic residues. Blagdon and Goodman¹³ suggested that these findings reflect interactions between the helix dipole and charged residues close to either terminus of an α -helix.

As noted earlier⁴, results from studies of block copolymers by Ooi and coworkers¹⁴ can also be interpreted in terms of a helix dipole model. In their system, helix formation by an (Ala)₂₀ block adjacent to a (Glu⁻)₂₀ block also shows a charged-group effect. For both these block copolymers and for the C-peptide derivatives studied here, the electrostatic effects are sequence-specific. Scheraga¹⁵ notes that sequence-specific interactions, which are not included in the Zimm-Bragg treatment of the helix-coil transition, are either missing or averaged out in the host-guest studies of random copolymers used to determine s and σ values.

Our results support Hol's view¹⁶⁻¹⁸ that interactions between charged groups and the α -helix macrodipole are functionally important in proteins. These interactions are likely to be even larger in proteins than in isolated α -helices in aqueous solution, because the effective dielectric constant in the interior of a protein is believed to be significantly lower than that of bulk water¹⁹. Recently, Perutz and co-workers²⁰ have discussed the possibility that a charged side chain interaction with a helix dipole is responsible for the high pK_a of a histidine residue in haemoglobin (see also ref. 21).

Another example comes from recent studies⁸ of semisynthetic ribonucleases formed by combination of S-protein (residues 21-124) with S-peptide derivatives (residues 1-15). These derivatives, based on the set of C-peptide analogues described here, contain varying NH₂-terminal charge. The results indicate a good correlation between stability of the isolated S-peptide helix and thermal stability of the reconstituted semisynthetic RNase S.

Possible mechanisms for a charged-group effect

By using a new reference peptide that shows good helix formation, we confirm, and demonstrate more decisively than before⁴, that Glu 2⁻ and His 12⁺ are the two charged residues responsible for the pH dependence of C-peptide helix content first noted

by Bierzynski *et al.*². This appears to be the major charged-group effect. As Fig. 2 indicates, however, there is residual helix in C-peptide analogues lacking either Glu 2 or His 12 at pH values for which the major charged-group effect is absent. Evidently, these electrostatic interactions do not account for the entire difference between the observed and predicted stabilities of the C-peptide helix⁴.

We introduced the helix dipole model to explain the charged-group effect assigned to Glu 2 and His 12 (ref. 4). In its simplest form, which can be used for making qualitative predictions, the helix dipole model approximates the actual distribution of partial charges in an α -helix as an extended line dipole with a positive pole near the NH₂-terminus of the helix and a negative pole near the COOH-terminus^{16,22}. Thus, electrostatic interactions between one pole of the helix macrodipole and a nearby charged group favour helix formation if the two charges are of opposite sign, and oppose helix formation if they are of like sign⁴.

The charged-group effect was first explained by postulating a Glu 9⁻ ... His 12⁺ ion pair², but later work ruled out this possibility⁴. Nevertheless, the effects of Glu 2⁻ and His 12⁺ might be explained by mechanisms other than the helix dipole model: both a Glu 2⁻ ... Arg 10⁺ ion pair and an aromatic interaction between Phe 8 and His 12⁺ have been suggested (ref 23, 29; K.R.S. *et al.*, unpublished data). Such side chain interactions are thought to be generally important for helix stabilization. The long central helices of troponin C and calmodulin may be stabilized by several intrahelical ion pairs and side chain hydrogen bonds²⁴. Glu⁻ ... Lys⁺ ion pairs contribute to the high stability of helices formed by recently designed model peptides seventeen residues in length (S. Marqusee and R.L.B., manuscript in preparation). Statistical analysis by Maxfield and Scheraga²⁵ indicates that the probability for a charged residue

to be helical is significantly greater if an amino acid of opposite charge is four residues away in the sequence.

In light of other postulated interactions in the C-peptide helix, it is important that we are able to test the helix dipole model here by varying the charge on the NH_2 -terminal residue. These studies (Figs 3 and 5) strongly suggest that interactions between a bound charged group and the helix dipole can be important. Screening of these charge-dipole interactions by mobile ions would explain the salt dependences of helix content shown in Fig. 4. More accurate models for making quantitative predictions of these effects will have to consider interactions between the bound charged groups and the point dipoles of individual peptide units. Such a modification of the Zimm-Bragg theory for the helix-coil transition has recently been introduced²⁶.

The charges that are approximated as monopoles in the line dipole model reside primarily at either end of the helix on the four NH and four CO groups that are not hydrogen-bonded. In their discussion of a sulphate-binding protein from *Escherichia coli*, Quijcho and Pflugrath²⁷ note that unsatisfied hydrogen-bonding donor groups at the NH_2 -termini of α -helices may account for the binding of negatively charged ligands close to the positive pole of the helix dipole. Two pieces of our data strongly suggest that the effects we observe are due to electrostatic interactions and cannot be explained by hydrogen-

bonding. Both screening of Lys^{12+} by NaCl (Fig. 4) and deprotonation of the $\alpha\text{-NH}_3^+$ group (Fig. 5a) appear to relieve repulsive interactions between bound positive charges and the positive pole of the helix dipole.

Short peptides and side-chain interactions

This work shows that the helix content of C-peptide can be almost doubled by combining a few favourable amino-acid replacements. Thereby, the phenomenon of α -helix formation by short peptides in aqueous solution is changed from a marginally observable experimental system to one that can be readily manipulated and studied. Thus, it is now practical to use short peptides to identify and analyse specific side-chain interactions that influence the stability of isolated α -helices in aqueous solution. Our results suggest that charged-group interactions with the helix dipole may stabilize early intermediates in protein folding.

We thank Virginia MacCosham for technical assistance, Dr S. W. Englander for discussion, Dr J. T. Yang for the use of his spectropolarimeter, and Virginia Sweeney for amino-acid analyses. This research was supported by the NIH (grant number GM 31475). K.R.S. is a predoctoral fellow of the NSF and P.S.K. was a predoctoral fellow of the Medical Scientist Training Program of the NIH.

Received 8 September 1986; accepted 6 February 1987.

1. Brown, J. E. & Klee, W. A. *Biochemistry* **10**, 470-476 (1971).
2. Bierzynski, A., Kim, P. S. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2470-2474 (1982).
3. Kim, P. S. & Baldwin, R. L. *Nature* **307**, 329-334 (1984).
4. Shoemaker, K. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2349-2353 (1985).
5. Zimm, B. H. & Bragg, J. K. *J. chem. Phys.* **31**, 526-535 (1959).
6. Sueki, M. *et al. Macromolecules* **17**, 148-155 (1984).
7. Holzwarth, G. & Doty, P. *J. Am. chem. Soc.* **87**, 218-228 (1965).
8. Mitchinson, C. & Baldwin, R. L. *Proteins* **1**, 23-33 (1986).
9. Baldwin, R. L. *Trends biochem. Sci.* **11**, 6-9 (1986).
10. Stewart, J. M. & Young, J. D. *Solid Phase Peptide Synthesis* (Pierce Chemical Company, Rockford, Illinois, 1984).
11. Wlodawer, A. & Sjölin, L. *Biochemistry* **22**, 2720-2728 (1983).
12. Chou, P. Y. & Fasman, G. D. *Biochemistry* **13**, 211-221 (1974).
13. Blagdon, D. E. & Goodman, M. *Biopolymers* **14**, 241-245 (1975).
14. Ihara, S., Ooi, T. & Takahashi, S. *Biopolymers* **21**, 131-145 (1982).
15. Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5585-5587 (1985).

16. Hol, W. G. J., van Duijnen, P. T. & Berendsen, H. J. C. *Nature* **273**, 443-446 (1978).
17. Hol, W. G. J., Halie, L. M. & Sander, C. *Nature* **294**, 532-536 (1981).
18. Hol, W. G. J. *Prog. Biophys. molec. Biol.* **45**, 149-195 (1985).
19. Matthew, J. B. A. *Rev. Biophys. Biophys. Chem.* **14**, 387-417 (1985).
20. Perutz, M. F., Gronenborn, A. M., Clore, G. M., Fogg, J. H. & Shi, D. T.-b. *J. molec. Biol.* **183**, 491-498 (1985).
21. Russu, I. M. & Ho, C. *Biochemistry* **25**, 1706-1716 (1986).
22. Sheridan, R. P., Levy, R. M. & Salemme, F. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4545-4549 (1982).
23. Rico, M. *et al. Biopolymers* **25**, 1031-1053 (1986).
24. Sundaralingam, M., Drendel, W. & Greaser, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7944-7947 (1985).
25. Maxfield, F. R. & Scheraga, H. A. *Macromolecules* **8**, 491-493 (1975).
26. Vasquez, M., Pincus, M. R. & Scheraga, H. A. *Biopolymers* **26**, 351-371 (1987).
27. Pflugrath, J. W. & Quijcho, F. A. *Nature* **314**, 257-260 (1985).
28. Chou, P. Y. & Fasman, G. D. *Adv. Enzym.* **47**, 45-148 (1978).
29. Bierzynski, A., Dadlez, M., Sobocinska, M. & Kupryszewski, G. *Biophys. Chem.* **25**, 127-134 (1986).

LETTERS TO NATURE

Pulsed TeV gamma-rays from Vela X-1

A. R. North, B. C. Raubenheimer, O. C. De Jager,
A. J. van Tonder & G. van Urk

PU-CSIR Cosmic Ray Research Unit, Potchefstroom University,
2520 Potchefstroom, South Africa

Emission of TeV γ -rays from the fast (periods less than 4 s) neutron stars in the X-ray binary systems Cyg X-3, Her X-1 and 4U0115+63 has recently been reported. Here we present the first evidence that the 283-s neutron star in the X-ray binary Vela X-1 radiates steady, pulsed TeV γ -rays with a sinusoidal light curve. This is also the first evidence that a wind-driven binary system radiates in this energy range. A periodic enhancement was also observed which was in phase with the steady pulsed emission and eight times stronger. This event was directly observable as an enhancement in the counting rate of the telescope and occurred while the pulsar was entering eclipse. This is considered to be an indication that the enhancement was caused by γ -ray production by a charged-particle beam interacting with the limb of the companion star. This fact further indicates that TeV γ -rays are not produced on the neutron star itself but in the accreting material in the system. This is probably the case with some of the other observed binaries.

The binary X-ray source Vela X-1 was discovered by the Uhuru and SAS 3 satellites¹ during 1975 and has been continuously monitored since then by a variety of satellites^{2,3}. The optical counterpart was identified⁴ as the B supergiant HD77581 at distance of (1.9 ± 0.2) kpc from the Earth. The orbital period is (8.9649 ± 0.002) days² whilst the pulse period of the neutron star is (282.84 ± 0.10) s. The X-ray emission as a function of orbital phase shows rapid and complicated variations as may be expected from a wind-driven, accreting binary^{1,3}. Large secular variations in the pulse period over timescales of days and years have been observed² but from 1980 to 1984 only a small spin-up ($\dot{P}/P = -2 \times 10^{-5} \text{ yr}^{-1}$) is evident³. The pulsed light curve is highly variable in the X-ray region changing from a very broad structure with five discernible peaks in the 5 keV range⁵ to a broad double peak in the (15-30) keV range^{2,6}. Vela X-1 was also reported to radiate in the PeV energy range^{7,8}.

Observations of Vela X-1 were made on eleven nights during the period from 2 April to 10 May 1986 with the TeV γ -ray telescope at Potchefstroom^{9,10}. This telescope is based on the atmospheric Cerenkov technique with a threshold energy of 1 TeV at the zenith and a field of view of 2.2° . On all occasions the source was tracked for a period of between 95 and 190 minutes at zenith angles between 17° and 36° , (observations at larger zenith angles were excluded due to the strong zenith angle dependence of the count rate). A total number of 54,450 Cerenkov events were registered during the observation time of 26.2 hours. The arrival time of each event was registered with a

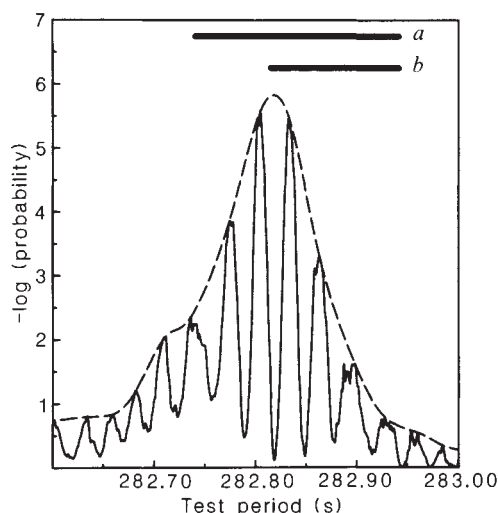


Fig. 1 The Rayleigh probability of chance origin of periodicity for the total data set of Vela X-1 observations as a function of the test period. The two main peaks define the most probable period range for periodic emission to be between 282.788 and 282.850 s. These values are within the bounds of all measurements of the pulse period³ made since its discovery (a) and are commensurate with the expected range for the duration of our measurements as extrapolated from satellite results³ between 1982 and 1984 (b). The dashed envelope is due to the 24-h modulation caused by the structure of the data.

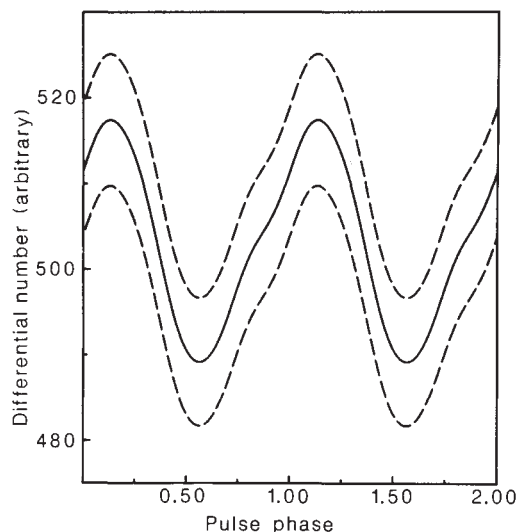


Fig. 2 The pulsed TeV light curve for Vela X-1 at 282.805 s. All the data covering integer pulsar periods were included and a normal kernel density estimator²⁰ was used to calculate the underlying light curve. The area under the light curve equals one hundredth of the total sample size. The dashed lines represent 95% confidence bands. The time of the first minimum is Julian Ephemeris day 2446523.2874 $\pm$ 0.0002.

resolution of 0.1 μ s. The data were corrected to the Solar System barycentre using the PEP740R ephemeris (J. F. Chandler *et al.*, personal communication). Due to the large velocities involved in this system the data were then transformed to the focus of the binary system using the most recent orbital parameters².

The most powerful test used to look for radiation at the fundamental harmonic of a period is the Rayleigh test¹¹. It was applied to the total data set and the search in period was confined to the range 282.74–282.94 s, the range over which the pulse period has varied over 12 years of observations². It is clear from Fig. 1 that there is a strong indication (a chance probability of 2×10^{-6}) of steady pulsed radiation in the assumed range. The narrow peaks are the independent peaks determined by the total duration of the observations (38 days). In uniformly distributed data the Rayleigh test statistic $2n\bar{R}^2$ will have a probability density function¹¹ proportional to $\exp(-0.5(2n\bar{R}^2))$. The large gaps in our data, as well as the relatively large periods tested, could influence this distribution and therefore the resulting chance probabilities. In fact this density was found in the data from a minimum χ^2 -fit at short periods (~ 28 s) but it changed to $(-0.4(2n\bar{R}^2))$ in the vicinity of the periods tested. This flatter distribution increased the above-mentioned chance probability to 2.8×10^{-5} . Although only eight independent periods were included in the defined period range, the number of effective independent trials due to oversampling in this interval was found from simulations to be 24. Thus, the probability for a chance origin of the emission at 282.805 s has to be increased to 6.6×10^{-4} . At this period the pulsed γ -ray content of the data was $3.1 \pm 0.6\%$ with all the power concentrated in the fundamental. Because the effective collection area of the telescope⁹ is $9 \times 10^8 \text{ cm}^2$ the resulting time-averaged pulsed flux during our observations is $(2.0 \pm 0.4) \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$. An analysis of the data on a night-by-night basis yields a standard deviation of 2.3% in the signal strength against the expected value of 2% if the samples were drawn from a distribution with a constant average signal strength¹¹. This indicates that the radiation was steady over the period of observation, implying no 8.9-day orbital modulation. Of the other known TeV γ -ray binaries only Cyg X-3 has a pronounced orbital modulation¹⁴. More measure-

ments on Vela X-1 are necessary to establish this because our observations only covered 10% of the binary orbit, scattered over the orbital phase range from 0.23 to 0.91.

In Fig. 2 the pulsed light curve of the data is illustrated. A broad and somewhat unsymmetric structure can be seen. This is not inconsistent with the X-ray results^{4,5} which indicate emission over a large part of the pulsar period. This light curve also shows some resemblance with the other TeV γ -ray binaries^{12–14}, Her X-1, 4U0115+63 and Cyg X-3. Assuming a mean threshold energy of 1.5 TeV for our telescope during these observations the γ -ray luminosity of Vela X-1, averaged over all our observations, is $(2.4 \pm 1.1) \times 10^{34} \text{ erg s}^{-1}$. This luminosity is about a factor of 10 lower than those measured so far for the above-mentioned X-ray binaries in the TeV range¹⁵. It must, however, be kept in mind that Vela X-1 is not only closer to Earth but, due to its longer orbital and pulsar period and the fact that it is wind driven, it is possible that the TeV γ -ray luminosity could be lower than those of the other systems.

Apart from the strong evidence for steady pulsed emission we also observed an outburst, which we are certain originates from the binary system. During the observation on 4 May 1986, lasting 123 min, the counting rate of all three independent telescope units in operation at that time increased for ~ 1.5 min by $\sim 100\%$ above the mean count rate at the start of that run (see Fig. 3). This is equal to a 5.7σ deviation from the mean rate with a probability of chance occurrence in our whole data set of 8.8×10^{-4} . Moreover, as may be seen from Fig. 3, it seems as if this enhancement is modulated with the pulsar period and that the enhancement persists over at least four pulsar rotations. This result, we believe, is the first direct illustration of periodic emission of TeV γ -rays from an X-ray binary. The non-pulsed excess for this event is equal to a flux of $(7.8 \pm 0.9) \times 10^{-10} \text{ cm}^{-2} \text{ s}^{-1}$. This enhancement is of the same order as that seen by various observers^{3,5,6} in the X-ray region. The whole run containing the outburst was then analysed for periodic emission. The data were split into 20 overlapping sections of length 4 times the pulsar period. The result indicated activity only during the time interval 8.8–27.6 min (see Fig. 3). During this time the probability that this emission at 282.805 s was due to chance is 6.5×10^{-4} , or less than 0.013 when the 20 overlapping trials were taken into account. Because the non-pulsed and

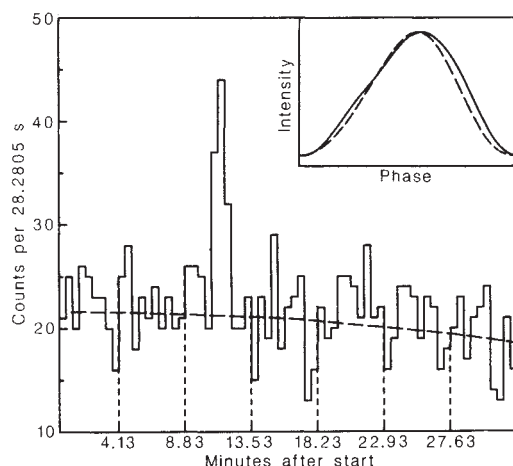


Fig. 3 The count rate of the observation on 4 May 1986 (zero time corresponds to 17:10 UT). The outburst started 11 min after the beginning of the run. The dashed curve represents a fit to the data excluding the outburst and is indicative of the zenith-angle dependence of the count rate. The vertical dashed lines correspond to the γ -ray minima, as calculated from Fig. 2. Inset: the pulsed light curve of the outburst (8.83 to 27.63 min) is represented by the dashed line. It is normalized in amplitude to the light curve of the steady radiation (solid line).

periodic tests on the outburst are independent, they may be combined, resulting in a probability of 1.1×10^{-5} that this pulsed outburst occurs by chance. A pulsed signal strength of $(18 \pm 5)\%$ is estimated for this event. This enhancement is in phase with the discussed steady emission and has the same light curve (see Fig. 3). Moreover, the pulsed signal strength deviates by 3.1σ from that of the steady emission so that this enhancement cannot be considered as a statistical fluctuation from the steady emission. It is important to notice that this run started at an orbital phase of 0.899, well within the entering stage of the eclipse¹⁶ which began four hours before the beginning of our observation, reaching full eclipse an hour after the start of the run. This event is reminiscent of the TeV observations of Her X-1 just after the start of eclipse¹⁷ and may support a recent model¹⁸ postulating γ -ray production by particle beams from the pulsar passing through the limb of the companion. This fact places Vela X-1 in the same category as Cyg X-3 and Her X-1 as far as the primary source of the γ -rays are concerned¹⁹, although Vela X-1 is different in that it is wind driven.

Concerning the PeV results on Vela X-1, the two reported claims for radiation^{7,8} are at orbital phases of 0.64 and 0.51 respectively. Inspection of these results did show less significant enhancements at phases around 0.15 and 0.84, near to eclipse. This may again be indicative of γ -ray production by charged particles in the accreting material or on the companion. It may be possible that PeV radiation from Vela X-1 also occurs during phases of enhanced TeV radiation. This is still conjecture and more extensive observations of the source in both the TeV and PeV regions are necessary. We may conclude, however, that the source of the TeV (and probably PeV) γ -rays are charged particles produced on the pulsar in the Vela X-1 system and that the TeV and PeV results do not contradict each other.

We are grateful to all our colleagues who offered their valuable time in realizing these measurements.

8. Suga, K. *et al.* in *Proc. Workshop on Techniques in UHE Gamma-Ray Astronomy*, La Jolla (eds Protheroe, R. J. & Stephens, S. A.), 48–53 (University of Adelaide, 1985).
9. De Jager, H. I. *et al.* *S. Afr. J. Phys.* **9**, 107–118 (1986).
10. Raubenheimer, B. C., North, A. R., De Jager, O. C., Van Urk, G. & Van Tonder, A. J. *Astrophys. J.* **307**, L43–L47 (1986).
11. Mardia, K. V. *Statistics of Directional Data* (Academic, London, 1972).
12. Dowthwaite, J. C. *et al.* *Nature* **309**, 691–693 (1984).
13. Chadwick, P. M. *et al.* *Astr. Astrophys.* **151**, L1–L3 (1985).
14. Watson, A. A. in *Proc. 19th Int. Cosmic Ray Conf.* La Jolla, Vol. 9 (ed. Jones, F. C.) 111–140 (NASA, Washington, 1985).
15. Weekes, T. C., *Proc. of the VIth Astrophys. Meet.*, Les Arcs, France (preprint, Center for Astrophysics, Cambridge, 1986).
16. Sato, N. *et al.* *Publ. astr. Soc. Japan* **38**, 723–730 (1986).
17. Gorham, P. W. *et al.* *Astrophys. J.* **308**, L11–L15 (1986).
18. Eichler, D. & Vestrand, W. T. *Nature* **318**, 345–347 (1985).
19. Ptuskin, V. S. *Proc. 19th Int. Cosmic Ray Conf.* La Jolla Vol. 9 (preprint, Center for Astrophysics, Cambridge, 1986).
20. De Jager, O. C., Swanepoel, J. W. H. & Raubenheimer, B. C. *Astr. Astrophys.* **170**, 187–196 (1986).

A massive glitch in an old pulsar

A. G. Lyne

University of Manchester, Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK

Pulsars normally spin down steadily due to the braking torque applied by the intense magnetic field of the neutron star¹. However, several of them display irregularities in rotation rate of two main types: random fluctuations, known as timing noise, and discrete discontinuities in rotation rate, known as glitches. Observations of these irregularities provide valuable information on the internal structure of neutron stars. Unfortunately glitches occur infrequently and large glitches have been observed in only two pulsars. Here, we describe a recent series of observations in which the relatively old pulsar PSR0355+54 has been seen to suffer two glitches, the second of which is the largest observed in any pulsar and provides a valuable opportunity for the study of the recovery of the interior of a neutron star following such a large perturbation.

PSR0355+54 was discovered² in 1972 and has been studied since then in a number of major timing programmes^{3,4}. The pulsar has a period (P) of 0.156 seconds, a period derivative (P') of $4.4 \times 10^{-15} \text{ ss}^{-1}$, and a characteristic age of $P/2P' = 0.6$ million years.

Timing observations have been made at Jodrell Bank since 1980 using the MkIA 76-m telescope, mostly at 408 MHz, but occasionally at 327 MHz, 610 MHz and 1,667 MHz. Post-detection de-dispersing receivers were used, allowing the total intensity to be recorded with a time resolution of better than one millisecond at all frequencies.

Pulse arrival times were obtained using the standard procedure of performing short online integrations synchronized to the apparent pulsar period and cross-correlating the resulting profiles with an appropriate template. These pulse arrival times were then reduced to the barycentre of the Solar System using the Jet Propulsion Laboratory planetary ephemeris DE200 (ref. 5) and the known position of the pulsar⁴. The results were combined with some of the measurements by Downs and Reichley⁴ spanning the years 1972–1980.

Until recently the timing measurements showed a very regular rotation, with a low level of timing noise and no glitches. The data between 1972.0 and 1985.0 are well fitted by a simple model consisting of the period and its first two derivatives (Fig. 1). Over this interval, the timing residuals from this model have an r.m.s. of only about 500 microseconds.

Figure 1 also indicates an abrupt decrease of period close to the beginning of 1985, amounting to a fractional change $\Delta P/P = -5.6 \times 10^{-9}$. Jumps of this magnitude or greater have been seen in about five other pulsars. The date of this glitch, glitch 1, is estimated as 14 January 1985 (modified Julian date (MJD) 46,079) ± 7 days, obtained by assuming that the glitch was simply an abrupt change of period, with no other short-term transient perturbations following the event.

Received 6 November 1986; accepted 5 February 1987.

1. Joss, P. C. & Rappaport, S. A. *A. Rev. Astr. Astrophys.* **22**, 537–592 (1984).
2. Nagase, F. *et al.* *Astrophys. J.* **280**, 259–268 (1984).
3. Nagase, F. *IAU Symp.* No. 125 (1986).
4. Sadakane, K. *et al.* *Astrophys. J.* **288**, 284–291 (1985).
5. Van der Klis, M. & Bonnet-Bidaud, J. M. *Adv. Space Res.* **3**, 39–42 (1984).
6. Bautz, M. *et al.* *Astrophys. J.* **266**, 794–805 (1983).
7. Protheroe, R. J., Clay, R. W. & Gerhardt, P. R. *Astrophys. J.* **280**, L47–L50 (1984).

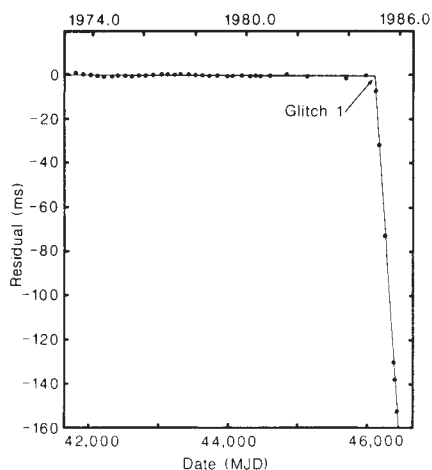


Fig. 1 The timing residuals for PSR0355+54 relative to a model fitted to observations before MJD 45962 (horizontal line). An abrupt change in period, glitch 1, is apparent about 100 days after.

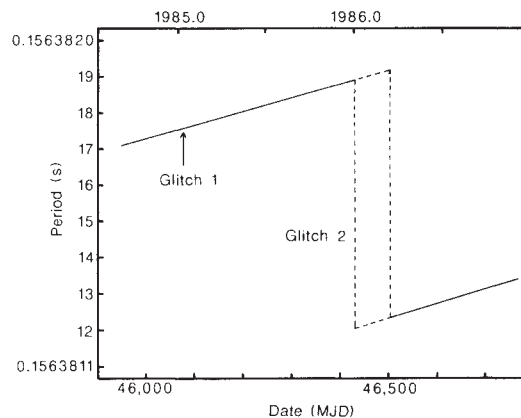


Fig. 2 The observed barycentric period of PSR0355+54 for the past 2 years. A massive glitch, glitch 2, interrupts the steady slow-down of the pulsar in early 1986. The broken lines show the earliest and latest paths the glitch could have taken. The earlier glitch 1 is not discernible on the scale of this diagram.

Table 1 Fitted parameters to the three data sequences

	Data span (MJD)	Epoch (MJD)	P (s)	error	P' (10^{-15})	error	P'' (10^{-24} s^{-1})	error
A	41807-45962	43,874.0	0.1563809205335	1	4.387778	2	-0.00291	3
B	46094-46433	46,250.0	0.156381820445	2	4.3916	2	-0.25	3
C	46504-46555	46,504.0	0.15638123391	2	4.66	2	-61	6
C'	46504-46773	46,504.0	0.15638123467	2	4.412	1	0	

The results are given for a multi-parameter fit to each of the data sequences A, B and C. The fit C' is for a model which includes an exponential term in the residual having a decay time of 44 ± 2 days and having an amplitude of 16 ± 1 ms at MJD 46504.0. Standard errors are given in units of the last quoted digit.

Then, on 15 March 1986, (MJD 46,504) a 25-min series of integrations showed the pulse rapidly drifting in phase, indicating that the period had increased drastically since the previous observation on 3 January 1986 (MJD 46,433). For this new event, glitch 2, the period change was $\Delta P/P = -4.4 \times 10^{-6}$, almost three orders of magnitude greater than for glitch 1. Figure 2 shows how the period of this pulsar has changed during the past two years. Glitch 1 is not discernible on the scale of this diagram. The period change associated with glitch 2 has reversed about five years of normal slow-down.

Unfortunately, the precise date of glitch 2 cannot be determined from these observations, since the pulse count is lost between the two adjoining measurements. The sheer magnitude of the glitch introduces an ambiguity equivalent to an extra rotation every 10 hours. It could have occurred any time between the observations of 3 January and 15 March.

The data also show that there were changes in the period first and second derivatives P' and P'' following the glitches. Fits for P , P' and P'' were made to observations before glitch 1, between the two glitches and immediately following glitch 2. Table 1 presents the three sets of best-fit parameters. (It should be remarked that the inclusion of P'' in the fits to the data does not imply that this quantity is a stationary deterministic value. The purpose of such fits is simply to provide a simple description of the rotation and its expected variation in the absence of glitches. In fact, it may be that many of the large values of P'' observed in other pulsars may arise, as in this case, in the recovery from glitches.)

Table 2 shows how these parameters changed at each of the two glitches. Because of the unknown date of glitch 2, the magnitude of the changes at the glitch are uncertain. The parameters given are the extreme values corresponding to the earliest and latest possible dates and are obtained by extrapolating the fits presented in Table 1. The slow-down rate increased by at least

6% and perhaps as much as 14%. This dramatic change is at least half as large again as the previous biggest $\Delta P'/P'$ for any pulsar.

A pulsar slows down as a result of the braking torque applied to the solid crust of the neutron star by the intense magnetic field. Under equilibrium slow-down conditions, part of the neutron superfluid in the neutron star may rotate faster than the solid crust, although it may be slowing down at the same rate. The departures from regular spin-down of a pulsar are then usually attributed to a variable coupling between the solid crust and this superfluid component. For instance, a sudden, brief increase in the coupling will result in a transfer of angular momentum from the superfluid to the solid crust and a reduction in the observed rotation period, followed by an increase in slow-down rate until the differential rotation redevelops.

In the vortex pinning model (for example, Pines and Alpar⁶) the steady slow-down of the superfluid involves the outward creep of quantized vortices which carry the rotation, and the angular momentum transfer is limited by the partial pinning of these rotational vortices to the crystalline lattice of the crust. The glitch results from a catastrophic unpinning of the vortices, which then move outwards from the rotation axis and repin; their outward movement reduces the density of vortices, that is, it reduces the rotation of the superfluid. The angular momentum

Table 2 The parameters of the two glitches

	Epoch (MJD)	Date	$\Delta P/P$ (10^{-9})	error	$\Delta P'/P'$	error
Glitch 1	$46,079 \pm 7$	85 Jan 14	-5.56	3	0.0018	2
	46,433	86 Jan 03	-4385	2	0.15	2
Glitch 2	↓	↓	↓	↓	↓	↓
	46,504	86 Mar 15	-4366.5	2	0.062	3

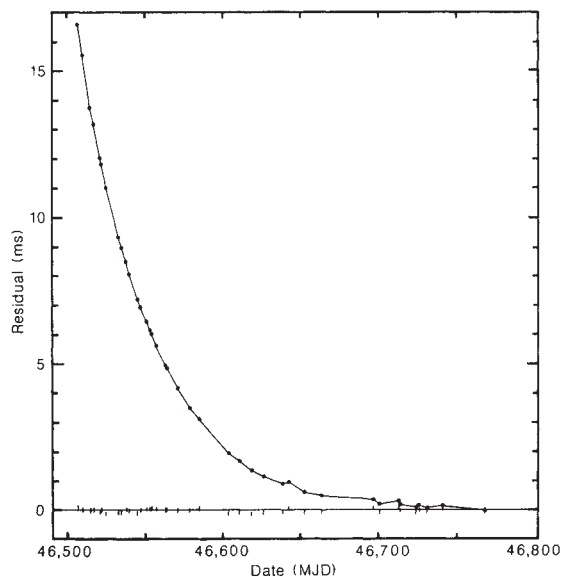


Fig. 3 The timing residuals for about eight months of data following glitch 2. The residuals shown by the vertical error bars near the bottom of the diagram are relative to a fitted model which uses the period, period derivative and an exponential term having a timescale of 44 days. The joined filled circles demonstrate the exponential nature of the recovery and show the residuals from this model without the exponential term.

of the superfluid is shared with the solid crust making it spin up.

This event results in a reduction in the difference of the rotation rate of the superfluid region through which the vortices moved during the glitch with respect to the rest of the crust. Immediately after the glitch, the vortices are completely pinned to the crystal lattice and vortex creep has stopped. The superfluid in the glitch region is then effectively decoupled from the rest of the star. This leads to a reduction in the effective moment of inertia and, because of the constant braking torque, a corresponding increase in P' at the glitch. Thus the quantity

$$\Delta P'/P' = \Delta I/I_c$$

indicates the fraction ΔI of the total moment of inertia of the crust I_c which is contained in the de-coupled superfluid⁷. The measured change in P' shows that $10 \pm 4\%$ of the pulsar's moment of inertia was involved in the large glitch 2. The good observational coverage following glitch 2 allows the recovery of the neutron star to be studied in some detail. Inspection of the residuals from the fit to these data shows that the simple P, P', P'' model does not give an adequate representation of the data. Indeed, the data are better described by a model which fitted only P and P' , together with an exponential decay in the residual or period having a characteristic timescale of 44 ± 2 days. Figure 3 shows the residuals with respect to this model for all the available data following glitch 2. The initial 44 day transient has decayed almost completely by the end of the observations as indicated by Fig. 3 although the period derivative is still greater than the pre-glitch value by an amount $\Delta P'/P' = 0.62 \pm 0.02\%$. This suggests that most of the de-coupled material has re-coupled again, although 0.6% of the moment of inertia is still de-coupled from the crust.

Because the torque on the pulsar is essentially constant during the glitch and subsequent recovery, by the time rotational equilibrium is re-established it is expected that the period will have recovered to the line described by the extrapolation of the period from the pre-glitch data in Fig. 2. Clearly, the period recovery so far has been very small. This recovery can only happen if the residual excess period derivative decays on a timescale of greater than $\Delta P/\Delta P' = 3 \times 10^{10} \text{ sec} = 1,000 \text{ yr}$.

It seems, therefore, that there may be two recovery timescales,

one of 44 days and one of at least a thousand years associated with components containing 10 and 0.6% of the total moment of inertia of the pulsar respectively. A similar two-component model, with shorter recovery timescales, is also required to account for the recovery from a glitch in the Vela pulsar^{8,9}. Another possibility is that the residual excess period derivative will decay on a much shorter timescale than 1,000 years, leaving a permanent change in the period as a result of the glitch. The simplex vortex pinning model does not allow for such a permanent change and an extra source of angular momentum for the spin-up is then required.

Because this pulsar is a very smooth rotator^{10,11}, and has had no other glitches since its discovery, it is likely that such glitches will occur rather infrequently. It should be possible to study the long-term recovery without perturbations from further glitches or timing noise as experienced in the Vela¹² and Crab¹³ pulsars. Whatever subsequent study reveals, both the magnitude of the glitch and the recovery demand some revision of accepted theories of the interiors of neutron stars. These observations also provide a caution that even though a pulsar may provide an excellent, dependable clock for many years, it can suffer a glitch without warning.

I thank Mr R. S. Pritchard for valuable assistance in the data analysis and Dr A. W. Jones and Professor F. G. Smith for constructive discussions.

Received 15 December 1986; accepted 3 February 1987.

1. Pacini, F. *Nature* **216**, 567-568 (1967).
2. Manchester, R. N., Taylor, J. H. & Huguenin, G. R. *Nature Phys. Sci.* **240**, 74 (1972).
3. Helfand, D. J., Taylor, J. H., Backus, P. R. & Cordes, J. M. *Astrophys. J.* **237**, 206-215 (1980).
4. Downs, G. S. & Reichley, P. E. *Astrophys. J. Suppl. Ser.* **53**, 169-240 (1983).
5. Standish, E. M., Keesey, M. S. W. & Newhall, X. X. *JPL Tech. Rep.* 132-160 (1976).
6. Pines, D. & Alpar, M. A. *Nature* **316**, 27-32 (1985).
7. Alpar, M. A. & Ho, C. *Mon. Not. R. astr. Soc.* **204**, 655-667 (1983).
8. McCulloch, P. M., Hamilton, P. A., Royle, G. W. R. & Manchester, R. N. *Nature* **302**, 319-320 (1983).
9. Alpar, M. A., Anderson, P. W., Pines, D. & Shaham, J. *Astrophys. J.* **278**, 791-805 (1984).
10. Cordes, J. M. & Downs, G. S. *Astrophys. J. Suppl. Ser.* **59**, 343-382 (1985).
11. Cordes, J. M. & Helfand, D. J. *Astrophys. J.* **239**, 640-650 (1980).
12. Downs, G. S. *Astrophys. J.* **249**, 687-697 (1981).
13. Demianski, M. & Proszynski, M. *Mon. Not. R. astr. Soc.* **202**, 437-451 (1983).

Ultraviolet detection of the nucleus of NGC2440

S. R. Heap

Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

In a recent letter, Atherton *et al.*¹ announced the detection of the central, exciting star of the planetary nebula, NGC2440, on a narrow-band charge-coupled device (CCD) image obtained under excellent seeing conditions (1 arc s). They found it to be one of the hottest stars known. Here, we report the detection of the same star in ultraviolet spectra taken with the International Ultraviolet Explorer (IUE) satellite². The stellar parameters derived from these spectra, $T \sim 200,000 \text{ K}$, $L = 1,500 L_{\odot}$, $\log g = 7$, $M = 0.6 M_{\odot}$, are in accord with stellar evolutionary theory.

Figure 1 compares the stellar flux distribution adopted by Atherton *et al.* (a black body at $T = 350,000 \text{ K}$ and $V_0 = 17.9$) with the underlying background flux computed (J. P. Harrington, personal communication) for the following nebular parameters³⁻⁵: $\log (F_{\beta})_0 = -10.07$, $T_e = 14,000 \text{ K}$, $N_e = 3,000 \text{ cm}^{-3}$, $N(\text{He}^+)/N(\text{H}) = 0.060$, $N(\text{He}^{+2})/N(\text{H}) = 0.053$. The average nebular flux per square arc s was then computed for a 200 square arc s emitting area and multiplied by the area of a disk having a diameter of 1, 2, and 5 arc s, the first two being representative of excellent and average seeing conditions, and the latter representative of the spatial resolution of the IUE spectrograph. For a given resolution, the star has its greatest contrast with respect to the nebula in the ultraviolet, particularly in the spectral

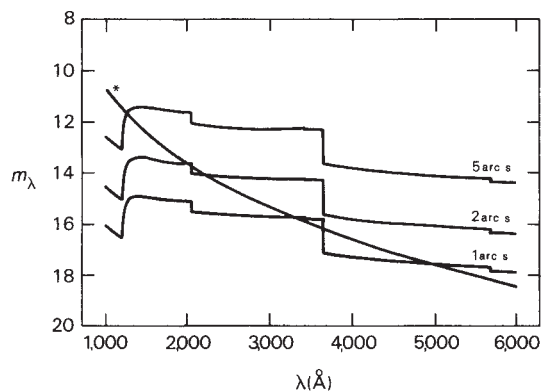


Fig. 1 Stellar and nebular flux distribution for NGC2440. The ordinate is monochromatic magnitude, defined as $m_\lambda = -2.5 \log F_\lambda - 21.1$. The nebular flux emitted from a disk 1, 2 and 5 arc s in diameter is also shown. The wavelength range of IUE spectrograms is roughly 1,150–3,250 Å, the wavelength of Altherton *et al.*'s picture is 4,788 Å.

region on the short wavelength side of Lyman- α (1,216 Å), where the only major source of nebular continuous emission is two-photon emission of He II. Even with the relatively coarse resolution of the IUE, the central star of NGC2440 should be evident in this spectral region.

On a photographic reconstruction of a long-exposure IUE spectrogram obtained with NGC2440 placed in the large (10×20 arc s) entrance aperture, a point-source continuum is visible at shorter wavelengths than that of Lyman- α down to the cutoff wavelength of the detector window (1,150 Å). In addition, a weak point-source continuum embedded in the extended nebular continuous emission appears at $\lambda < 1,400$ Å. We interpret this point-source emission as due to the central star.

Figure 2 shows a decomposition of the observed spectrum into its stellar and nebular components. The observed spectrum was first corrected for interstellar extinction⁶ with $E_{B-V} = 0.30$. Regions free of nebular line emission were binned (usually over 30 Å) to obtain a measure of the observed continuum. The calculated nebular continuum was multiplied by a factor, $A = 0.7$, the fraction of nebular light subtended by the spectrograph aperture, in order to match the observed flux at longer wavelengths ($\lambda > 1,900$ Å) where the stellar contribution is negligible. The star was assumed to be a black body at 200,000 K, whose flux was scaled so that when added to the theoretical nebular spectrum, the sum would match the observed flux distribution at the shorter wavelengths ($\lambda < 1,900$ Å). The best fit, shown in Fig. 2, corresponds to an unreddened stellar magnitude of $V_0 = 16.8$ at 5,480 Å.

To rephrase the preceding paragraph, an E_{B-V} of 0.30 yields a flux distribution implying a stellar visual luminosity three times that estimated by Atherton *et al.* If, however, E_{B-V} were reduced to 0.20, then the discrepancy between the results of Atherton *et al.* and ours would be eliminated. The precise value of E_{B-V} is therefore crucial. Fortunately, its value is sufficiently well determined. The adopted value³, $E_{B-V} = 0.30 \pm 0.03$, is based on the ratio of radio to H_β flux from the nebula, as this is generally considered to be the best measure of interstellar extinction. The He II $\lambda 1,640/\lambda 3,203$ nebular flux ratio obtained from IUE spectra also gives⁷ $E_{B-V} = 0.35 \pm 0.04$ and thus confirms that the unreddened ultraviolet flux distribution in Fig. 2 is basically correct. Estimates of extinction based on the Balmer decrement are also on average consistent, although there is an unaccountably wide variation from author to author⁴.

Line-flux ratios that combine IUE measurements, which sample only a portion of the nebula, with ground-based observations that include the whole nebula then yield values of A consistent with the assumed colour excess, $E_{B-V} = 0.30$. The He II $\lambda 1,640/4,686$ flux-ratio yields⁷ $A = 0.53$, and the [O III]

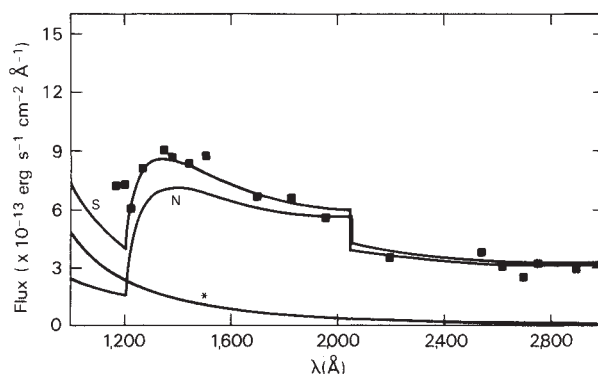


Fig. 2 Ultraviolet spectrum of NGC2440 derived from IUE observations, (SWP) 14,108 and (LWR) 10,074, corrected for extinction with $E_{B-V} = 0.30$. The contributions of the star (*), nebula (N), and sum (S) are also shown.

$\lambda 1,663/4,959$; 5007 ratio gives⁵ $A = 0.63$. As mentioned before, the continuum flux distribution yields $A = 0.7$. We retain $A = 0.53$ and $A = 0.70$ as possibilities.

We next estimate the temperature of the central star using the Zanstra method⁸ applied to He II line at 1,640 Å. We measured the nebular line-emission in five IUE spectra in which the line at 1,640 Å was not saturated to obtain 2.2×10^{-11} erg s⁻¹ cm⁻², which corrected for interstellar extinction gives 2.0×10^{-10} erg s⁻¹ cm⁻². The corrected stellar monochromatic flux in the region of the nebular lines is 7.7×10^{-14} erg s⁻¹ cm⁻² Å⁻¹. These quantities yield the product, $AQ = 3.1$, where Q is a temperature-sensitive parameter. The derived black-body temperature is about 190,000 K for $A = 0.70$ (continuum estimate)

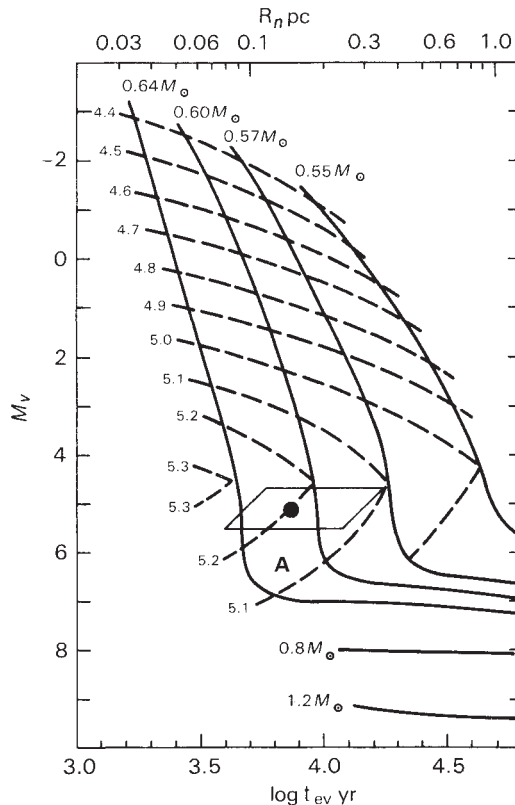


Fig. 3 Evolutionary diagram for planetary nuclei (ref. 12). Solid lines show evolutionary tracks for stars of the indicated masses; dashed lines shown contours of log temperature. The location of NGC2440 as estimated by Atherton *et al.*¹ (A) and this study (●) is indicated. The error box includes uncertainties in distance and angular radius of the nebula.

and 210,000 K for $A = 0.53$ (He II lines). These estimates are in agreement with $T = 170,000$ – $230,000$ K, obtained by the energy-balance method^{1,9}. We therefore adopt $T = 200,000$ K.

Given a distance³ to NGC2440 of $2,190 \pm 400$ pc (corresponding to $E_{B-V} = 0.30 \pm 0.03$), the stellar absolute magnitude is $M_V = +5.05 \pm 0.4$. The basic stellar parameters corresponding to $T = 200,000$ and $M_V = +5.05$ ($L = 1,500 L_\odot$, $R = 0.03 R_\odot$, and $\log g = 7$) indicate that the star is a hot white dwarf. Its age since ejecting the nebular is 7,100 yr, based on a nebular radius⁵ of 15 arc s and an expansion velocity¹⁰ of 22 km s^{-1} .

Are these characteristics in accord with stellar evolutionary theory? The theory^{10,11} predicts that a central star of a planetary nebula evolves from a red giant, crosses the H-R (Hertzsprung–Russell) diagram from right to left at a constant luminosity, and then descends along the white-dwarf cooling track for its mass. During its horizontal traverse, the star grows hotter, and a greater fraction of its luminosity is radiated in the extreme ultraviolet, a lesser fraction in the optical region of the spectrum. Thus, the star appears to fade with time. The rate of fading is exceedingly dependent on stellar mass; hence, the position of the nucleus of NGC2440 on an evolutionary diagram¹² (Fig. 3) is a sensitive mass-indicator. The inferred mass, $0.62 M_\odot$, is in keeping with its stellar population characteristics and is typical of planetary nuclei as a class^{12,13}. The position of the nucleus of NGC2440 on the diagram also implies a stellar temperature of about 160,000 K, in coarse agreement with temperature estimates

based on the Zanstra method or the method of energy-balance. We therefore conclude that the observed characteristics of NGC2440 are in accord with stellar evolutionary theory.

In contrast, Atherton *et al.* derived different stellar parameters ($T = 350,000$ K; mass = $1.0 M_\odot$) which conflict with predictions, and they concluded that “the theory is incomplete”. Although their conclusion may prove to be correct, we argue that the case of NGC2440 does not necessarily support their case. Their apparent discrepancies can be resolved if their estimated apparent magnitude of the star is raised from $V = 18.9$ to 17.8. In view of the extreme difficulty of even detecting the star, any estimate of apparent magnitude should presently be regarded as tentative and await definitive measurement with the Hubble Space Telescope.

Received 15 May 1986; accepted 15 January 1987.

1. Atherton, P. D., Reay, N. K. & Pottasch, S. R. *Nature* **320**, 423–425 (1986).
2. Boggess, A. *et al.* *Nature* **275**, 1–15 (1978).
3. Gathier, R., Pottasch, S. R. & Pel, J. W. *Astr. Astrophys.* **151**, 171–190 (1986).
4. Gutierrez-Moreno, A., Moreno, M. & Cortes, G. *Publ. astr. Soc. Pacif.* **97**, 397–403 (1985).
5. Shields, G. A., Aller, L. H., Keyes, C. D. & Czyzak, S. J. *Astrophys. J.* **248**, 569–583 (1981).
6. Seaton, M. J. *Mon. Not. R. astr. Soc.* **187**, 73P–76P (1979).
7. Seaton, M. J. *Mon. Not. R. astr. Soc.* **185**, 5P–8P (1978).
8. Bohlin, R., Harrington, J. P. & Stecher, T. P. *Astrophys. J.* **252**, 635–643 (1982).
9. Preite-Martinez, A. & Pottasch, S. R. *Astr. Astrophys.* **126**, 31–44 (1983).
10. Wilson, O. C. *Astrophys. J.* **111**, 279–305 (1950).
11. Paczyński, B. *Acta astr.* **21**, 417–435 (1971).
12. Schönberner, D. *Astr. Astrophys.* **103**, 119–130 (1981).
13. Heap, S. R. & Augenson, H. *Astrophys. J.* **313**, 268–283 (1987).

Rejection of a proposed 7.4-day rotation period of the comet Halley nucleus

B. A. Smith*, S. M. Larson*, K. Szegö†
& R. Z. Sagdeev‡

* Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

† Central Research Institute for Physics, PO Box 49, H-1525 Budapest 114, Hungary

‡ Space Research Institute, Profsoyuznaya 84/32, 117810 Moscow GSP-7, USSR

Preliminary analyses of Vega 1 and Vega 2 images of the comet Halley nucleus have led to a derived rotation period of 53.5 ± 1 hours (2.23 ± 0.04 days) about an axis approximately perpendicular to the long axis of the nucleus^{1–3}. Images obtained by Giotto^{4–6} were also found to be consistent with the Vega-derived rotation period. Recently, however, a longer rotation period of 7.4 days has been proposed⁷, based on an apparent double-peaked periodicity in the production rates of CN, C₂ and dust. Although we acknowledge possible photometric evidence for time-variable production rates and appreciate the assumption that this periodicity should somehow be related to the rotation of the nucleus, we conclude, nevertheless, that the imaging observations from the three spacecraft provide a compelling refutation of the proposed 7.4-day rotation period. Of crucial importance to this conclusion are: (1) the non-symmetrical shape of the Halley nucleus; and (2) the independent observations of this irregular body by Vega 1, Vega 2 and Giotto over an interval of 7.7 days. We note, however, that a rotation period of 53.5 hours and a longer period of torque-free precession with the angular momentum vector inclined somewhat to the spin axis might be consistent with both the spacecraft and ground-based observations.

The approximate shape of the Halley nucleus, as revealed by the Vega and Giotto cameras, is that of a triaxial ellipsoid measuring $\sim 7.5 \times 8 \times 16$ km; in detail, however, the actual shape departs markedly from an ellipsoid. As can be seen in Fig. 1, the nucleus is quite obviously larger at the left end, 7.5 km across



Fig. 1 Vega 2 image of the comet Halley nucleus taken 1.5 s before closest approach. The larger end of the nucleus (left) is directed away from the camera at an angle of about 45° to the image plane. The Sun is to the left. North is at the top.

as opposed to a 5-km width at the right end. It also exhibits a distinctive north-south asymmetry at the larger end and has a one kilometre deep depression near the centre of its southern face. It is this non-symmetrical shape that makes the spatial orientation of the nucleus easily recognizable in those Vega and Giotto images which have adequate spatial resolution.

Figure 2 shows the observed orientation of the long axis of the nucleus, projected onto the plane of the ecliptic, with respect to the direction of the Sun at the moments of encounter by Vega 1, Vega 2 and Giotto; the viewing direction of critical observations by the three spacecraft is also indicated. Let us now consider that the nucleus does indeed rotate about an axis nearly perpendicular to its long dimension with a period of 7.4 days. Although this is the most stable spin configuration, it is not the only one allowed and other spin states will be examined later. Because the interval between the Vega 1 and Giotto encounters was 7.7 days, the proposed 7.4-day period implies that the nucleus made slightly more than a single rotation between the two encounters; thus, the same end of the nucleus

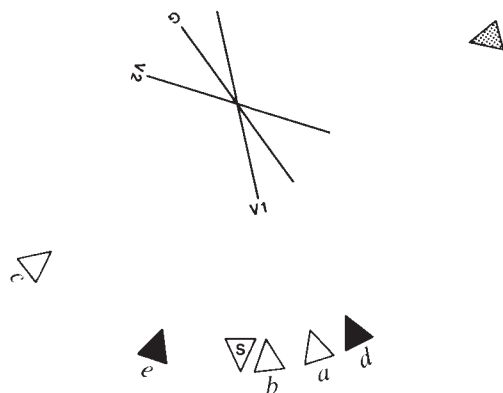


Fig. 2 Critical viewing directions relative to the long axis of the nucleus and the direction (S) of the Sun. Vega 1 (V1), empty triangles; Vega 2 (V2), solid triangles; Giotto (G), stippled triangle. a, Encounter (E) + 12.4 s; b, E + 34.4 s; c, E + 383.4 s; d, E - 1.5 s; e, E + 98.7 s. The solid lines show the projection onto the ecliptic plane of the long axis of the nucleus.

must have faced the Sun during both Vega 1 and Giotto encounters. We also note that the nucleus could have rotated only 146° during the 3.0-day interval between Vega encounters. This implies that only direct rotation (clockwise in Fig. 2) is satisfied and, because the Vega 2 image (Fig. 1) clearly shows the large end to the left, the same large end must have been pointing toward the Sun at the time of the Vega 1 encounter. The 7.4-day period, therefore, also requires that it must have been the large end that was pointed more or less towards the Sun at the time of the Giotto encounter. The evidence, however, is to the contrary.

Although Reitsema *et al.*⁶ have claimed that the large end of the Halley nucleus was pointed into the anti-solar hemisphere at the time of Giotto encounter, proponents of the 7.4-day rotation period have warned that dust obscuration around the illuminated end makes such an identification questionable. We will now use the non-symmetrical shape of the nucleus to support the conclusion of Reitsema *et al.*⁶. Figure 3 shows profiles taken from the Vega 2 image seen in Fig. 1 (corrected for foreshortening) and one of the many Giotto images of the nucleus, all of which were seen silhouetted against the coma^{4,6}. Because the Giotto profile in Fig. 3 is a left-to-right mirror image of the real view, we can conclude that opposite sides of the nucleus were being seen by the Vega 2 and Giotto spacecraft. The mirror reversal of the Giotto image also places the Sun to the right, and so, from the remarkably good matching of this irregular profile, it is evident that it was the small end of the nucleus that was actually directed toward the Sun during the Giotto flyby. Thus, a rotation period of 7.4 days about an axis approximately perpendicular to the long axis of the nucleus must be rejected.

In an attempt to resolve the dilemma created by the spacecraft optical observations and the ground-based photometry, Sekanina⁸ has proposed that the 7.4-day rotation actually takes place about the long axis of the nucleus, and that torque-free precession occurs with a 2.2-day period at a precession angle of 77°. We examine this possibility to see if it is consistent with the spacecraft observations. As it turns out, the Sekanina proposal is quickly dealt with for any probable geometrical shape of the Halley nucleus; the non-symmetrical profile of the large end, shown in two views in Fig. 3 obtained 4.7 days apart, can allow only an integer number of rotations about the long axis over an interval of 4.7 days. Thus, a 7.4-day rotation period about the long axis must also be rejected.

How then can we explain the observed 7.4-day, double-peaked period for gas and dust production? Julian⁹ has proposed that a body having the approximate shape of the Halley nucleus can undergo torque-free precession with a precessional period that could explain the photometric observations. If such a precession were impressed upon a 2.2-day rotation about an axis nearly

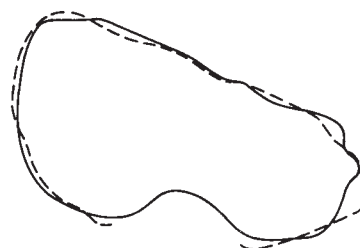


Fig. 3 Profiles of the comet Halley nucleus seen by Vega 2 (solid line) and Giotto (dashed line). The Vega profile has been corrected for foreshortening; the small end is partially in shadow. The Giotto profile is a mirror image, indicating that Giotto viewed the opposite side of the nucleus to that seen by Vega 2; the limb along the bottom is indefinite because of strong jet activity. The approximate direction of the Sun at the time of the Giotto encounter is towards the right in the mirror-image profile.

normal to the long axis, then highly active regions on the surface of the nucleus could be exposed to occasionally high solar elevation angles at intervals that might explain the observed periodicity in gas and dust production.

There is, in fact, some observational support for free precession. The inclination of the long axis of the nucleus to the ecliptic plane as seen at different times by the three spacecraft does not appear to be consistent with rotation axis which is fixed with respect to a non-rotating coordinate system centred on the centre of mass of the nucleus. The observed set of inclinations is perhaps more consistent with a small precession angle of at most a few tens of degrees. A more detailed analysis is now being carried out jointly by members of the Vega and Giotto imaging teams.

In summary, the 7.4-day period cannot be taken as the rotation period of the Halley nucleus, although torque-free precession might explain the observed periodicity in gas and dust production rates. We stand by our previously reported period of 53.5 h.

Received 2 December 1986; accepted 11 February 1987.

1. Sagdeev, R. Z. *et al. Nature* **321**, 262-266 (1986).
2. Sagdeev, R. Z. *et al. COSPAR Proceedings*, Toulouse (in the press).
3. Sagdeev, R. Z. *et al. Proc. 20th ESLAB Symp. on the Exploration of Halley's Comet*, Vol. II (eds Battrick, B., Rolfe, E. J. & Reinhard, R.) 335-338 (ESA Publications Division, ESTEC, Noordwijk, The Netherlands, 1987).
4. Keller, H. U. *et al. Nature* **321**, 320-326 (1986).
5. Keller, H. U. *et al. Proc. 20th ESLAB Symp. on the Exploration of Halley's Comet*, Vol. II (eds Battrick, B., Rolfe, E. J. & Reinhard, R.) 347-350 (ESA Publications Division, ESTEC, Noordwijk, The Netherlands, 1987).
6. Reitsema, H. J. *Proc. 20th ESLAB Symp. on the Exploration of Halley's Comet*, Vol. II (eds Battrick, B., Rolfe, E. J. & Reinhard, R.) 351-354 (ESA Publications Division, ESTEC, Noordwijk, The Netherlands, 1987).
7. Millis, R. L. & Schleicher, D. G. *Nature* **324**, 646-649 (1986).
8. Sekanina, Z. *Nature* **325**, 326-328 (1987).
9. Julian, W. H. *Nature* **326**, 57-58 (1987).

Oscillations in photon distribution of squeezed states and interference in phase space

W. Schleich* & J. A. Wheeler

Center for Theoretical Physics, University of Texas at Austin, Austin, Texas 78712, USA

The drive for both noise-free message transmission^{1,2} and high precision gravity wave detection^{3,4} has stimulated immense effort on a key element, a squeezed state^{5,6} of the electromagnetic field. Such non-classical states have been investigated theoretically in great detail¹⁻⁷ and have now been realized experimentally in four laboratories in the United States⁸⁻¹³. However, nowhere in the literature have we been able to find the striking feature of a

* Present address: Max-Planck-Institut für Quantenoptik, D-8046 Garching b. München, FRG.

squeezed state which we report here: an oscillatory distribution in photon number^{14,15}. These oscillations, and the conditions which produce them, came to light in the course of an investigation of sudden transitions¹⁶ (the Franck–Condon effect in molecular physics^{17,18}) based on the semi-classical description of a quantum state¹⁹ as motion of a representative point in the phase space defined by oscillator coordinate and oscillator momentum.

The essential feature of a squeezed state of the field can be translated into the language of a harmonic oscillator of mass μ and frequency ω . A massive cylinder rolling under the influence of gravity on a metal ruler bent into the shape of a parabola can serve as a model. To begin with, the oscillator is in its lowest quantum state¹⁶

$$\psi_0(q) = \left(\frac{\mu\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\frac{\mu\omega}{2\hbar}q^2\right\} \quad (1)$$

The term 'squeezed state' in current usage denotes a state obtained from this ground state by a combination of simultaneous sudden 'squeeze' and sudden displacement. In the mechanical model this 'squeezing' is achieved by a sudden alteration of the steepness of the potential with a corresponding change of the frequency from ω to $\omega' = s^{-1}\omega$ where $s > 0$. According to equation (1) the width Δq of the probability distribution $|\psi_0|^2$ is $\Delta q = s^{-1/2}(\hbar/\mu\omega')^{1/2}$. Thus, for $s > 1$ we have reduced, or squeezed the uncertainty in position q , whereas for $0 < s < 1$ this uncertainty is enhanced. In addition, we suddenly displace the origin of the harmonic potential by an amount $q_0 = (2\hbar/\mu\omega')^{1/2}\alpha$ and lower the potential energy by $(1/2)\mu\omega'^2 q_0^2 = \hbar\omega'\alpha^2$.

We use the dimensionless coordinate $x = (\mu\omega'/\hbar)^{1/2}q$. The wavefunction of the state, so prepared—a squeezed state—reads

$$\psi_{sq}(x) = \pi^{-1/4} s^{1/4} \exp\left\{-\frac{1}{2}s(x - \sqrt{2}\alpha)^2\right\} \quad (2)$$

The corresponding distribution of probability in (dimensionless) momentum p is given by

$$\phi_{sq}(p) = \pi^{-1/4} s^{-1/4} \exp\left\{-\frac{p^2}{2s}\right\} \quad (3)$$

Equations (2) and (3) indicate clearly that squeezing the uncertainty $\Delta x = s^{-1/2}$ in x is achieved at the expense of the momentum uncertainty $\Delta p = s^{1/2}$ so that the minimum uncertainty relation $\Delta x \cdot \Delta p = 1$ is maintained.

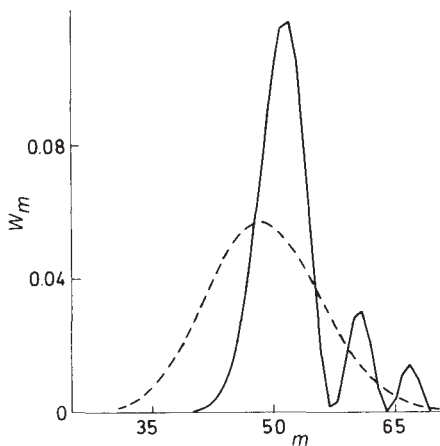


Fig. 1 Probability W_m of finding m photons in a coherent state (broken line) or in a highly squeezed state (solid line), based on equation (6). (Squeezing parameter $s = 21$ and displacement $\alpha = 7$. Every classical oscillator initially at rest and then subjected to such a displacement will undergo the same sudden increase in energy, an increase which in our units is $\alpha^2 = 49$). Curves, it should be recognized, are not curves, because m is never other than an integer.

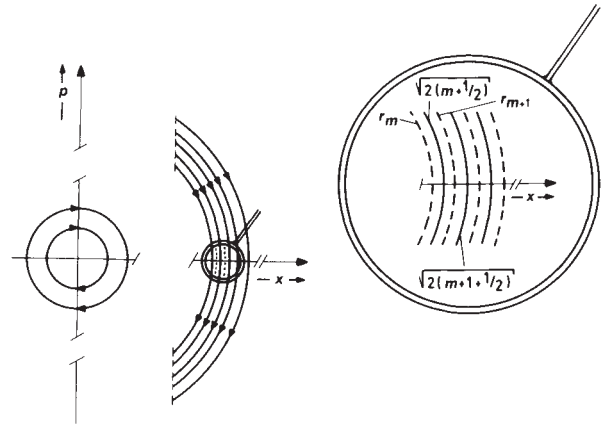


Fig. 2 A single mode of the electromagnetic field in a number state is equivalent to a harmonic oscillator with dimensionless coordinates x and p . The trajectories in phase space are circles with radius $\sqrt{2(m+1/2)}$ traversed in the clockwise direction. With each state we associate a band of area 2π (in units $\hbar$) defined by its inner radius $r_m = \sqrt{2m}$ and its outer radius $r_{m+1} = \sqrt{2(m+1)}$.

Both squeezing and displacement have altered the energy of the oscillator. The squeezing has changed the ground-state energy, expressed in these dimensionless units, from $1/2$ to $(1/2)s$. Moreover, the displacement has introduced the energy α^2 . However, the energy of the oscillator in a squeezed state is not the sum of the two energies. There exists no deterministic formula for its energy. Instead there is a probability W_m that the oscillator is in its m th energy eigenstate

$$u_m(x) = \pi^{-1/4} (2^m m!)^{-1/2} H_m(x) e^{-(1/2)sx^2} \quad (4)$$

Here H_m denotes the m th Hermite polynomial²⁰. According to the standard rules of quantum mechanics¹⁶ W_m is given by

$$W_m = \left| \int_{-\infty}^{\infty} dx u_m(x) \psi_{sq}(x) \right|^2 \quad (5)$$

In the language of quantized electromagnetic fields W_m represents the probability of finding m photons in a squeezed state.

We substitute equations (2) and (4) into equation (5), perform the integration and arrive^{1,2,7} for $s \geq 1$ at

$$W_m = \frac{2\sqrt{s}}{s+1} \left(\frac{s-1}{s+1}\right)^m (2^m m!)^{-1} H_m^2\left(\frac{s}{(s^2-1)^{1/2}}\sqrt{2}\alpha\right) \times \exp\left\{-\frac{2s}{s+1}\alpha^2\right\} \quad (6)$$

In Fig. 1 we compare the photon distribution W_m obtained for $s = 1$ (broken line) to the corresponding one for $s = 21$ (solid line). In both cases we adopt the value $\alpha = 7$ for the displacement parameter. For the case $s = 1$ we know that $\Delta x = \Delta p = 1$, that is, the fluctuations are equally distributed on the two quadratures of the field. Such coherent states are known to play a prominent role in quantum optics²¹. On the other hand the case $s = 21$ yields $\Delta x \approx 0.2$ and $\Delta p \approx 5$ and therefore corresponds to a highly squeezed state.

Whereas the coherent state obeys the familiar Poisson distribution²¹ shown by the broken line, the squeezed state exhibits oscillations for photon numbers $m > \alpha^2$. These modulations, hidden in equation (6) within the special functions, come to light in the asymptotic expansion of W_m (J.A.W. and W.S., manuscript in preparation)

$$W_m = 4\mathcal{A}_m \cos^2 \phi_m \quad (7)$$

Equation (7) is valid for large squeezing ($s = 2/\epsilon$, where $0 < \epsilon \ll 1$) and large quantum numbers $m > \alpha^2 \gg 1$. According to

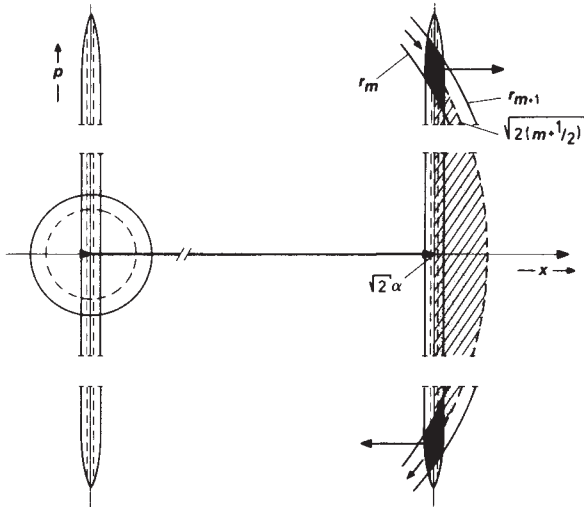


Fig. 3 A squeezed state is obtained by first 'squeezing' the ground state in one (here x -) variable, at the expense of the other (p -) variable and then displacing it by $\sqrt{2}\alpha$. For strong squeezing the contour lines in phase space (lines where P_{sq} is constant), have the shape of a long, thin ellipse ("Gaussian cigar"). Therefore, a band defined by the radii r_m and r_{m+1} with m sufficiently larger than α^2 "intersects" this cigar twice in the two symmetrically located dark areas. In one zone the oscillator is moving to the right, in the other, to the left. The total probability amplitude, $\sqrt{W_m}$ is the sum of contributions $\sqrt{\mathcal{A}_m} \exp(\pm i\phi_m)$ from the dark areas. The phase ϕ_m (equation (9)) is fixed by the shaded area. No simpler illustration is given for "interference in phase space".

equation (7) the decay of W_m associated with the amplitude factor $\mathcal{A}_m$

$$\mathcal{A}_m = \left(\frac{\varepsilon}{4\pi}\right)^{1/2} \frac{\exp\{\varepsilon(\alpha^2 - m - 1/2)\}}{(m + 1/2 - \alpha^2)^{1/2}} \quad (8)$$

is modulated with the phase

$$\phi_m = \frac{1}{2} \left(\int_{-\sqrt{2}\alpha}^{\xi_m} dx \int_{-p_m}^{p_m} dp - \frac{\pi}{2} \right) \quad (9)$$

where $\xi_m = \sqrt{2(m+1/2)}$ and $p_m = \sqrt{2(m+1/2) - x^2}$. This modulation, however, can only be noticed if the decay length of $\mathcal{A}_m$, $1/\varepsilon$ is of the order of, or larger than, the period $\Delta m \sim \alpha^{2/3}$ of the oscillations following from equation (9) in the limit of large displacements (J.A.W. and S.W., manuscript in preparation). Since $s = 2/\varepsilon$ we arrive at the following condition for oscillations,

$$s \gtrsim \alpha^{2/3}$$

This condition ensures the existence of the oscillations. But why oscillations? What is the underlying physics? The answer is interference. Interference is as inescapable in the physics of the squeezed state as it is in Young's experiment^{22,23}. Here, however, the interference takes place in a more sophisticated arena: not in coordinate space, but in phase space, as we show in the remainder of this report.

In the semi-classical limit (Bohr's correspondence principle^{19,24,25}) the transition probability W_m (equation (5)) between two quantum-mechanical states such as a number state and a squeezed state is governed¹⁴ by the area of overlap between the two states represented as bands in phase space.

The energy in the m th number state reads

$$m + \frac{1}{2} = \frac{1}{2} p_m^2 + \frac{1}{2} x^2 \quad (10)$$

Therefore the trajectories in phase space are circles of radius $\sqrt{2(m+1/2)}$ traversed in a clockwise direction. Each state takes up an area 2π in phase space²⁶. Therefore, we associate with the m th number state an occupied band of inner radius $\sqrt{2m}$

and outer radius $\sqrt{2(m+1)}$ as shown in Fig. 2.

The ground state ($m=0$), in particular, is represented by a circle of radius $\sqrt{2}$ and its centre curve by a circle of radius unity as indicated in Fig. 3 by a solid and broken line, respectively. Strongly squeezing this state in the x -variable at the expense of the momentum and displacing it by an amount of $\sqrt{2}\alpha$ results in a long thin cigar, vertical and centred at the point $x_0 = \sqrt{2}\alpha$ on the positive x -axis.

According to the area-of-overlap concept¹⁴, the probability W_m of finding m photons in the squeezed state is governed by the overlap in phase space between the circular band representing the m th number state and the tall cigar representing the squeezed state. The cigar and the ring overlap either not at all or in two diamond-shaped areas as indicated in Fig. 3, according to whether the excitation m is appropriately greater or less than α^2 . The area of overlap A_m associated with either one of these diamonds we obtain by weighting each point of phase space according to the distribution function P_{sq} ^{27,28}

$$P_{sq}(x, p) = |\psi_{sq}(x)|^2 |\phi_{sq}(p)|^2 = \pi^{-1} \exp \left\{ -\frac{2}{\varepsilon} (x - \sqrt{2}\alpha)^2 - \frac{\varepsilon}{2} p^2 \right\}$$

which expresses the internal structure of the "Gaussian cigar" given by equations (2) and (3). The area of overlap A_m is thus

$$A_m = \frac{1}{2} \int_{m\text{-th band}} dx \int dp P_{sq}(x, p) = \mathcal{A}_m;$$

that is, the area of overlap is identical to the quantity defined by equation (8).

The probability to be compared with experiment is not the sum, $2A_m$ of the areas of the two diamonds. Neither is the intensity on the photographic plate in the familiar double-slit experiment equal to the sum of the intensities that would arrive through the two slits separately!

Quantum mechanics instructs us to add, not probabilities, but probability amplitudes. The absolute value of the probability amplitude corresponding to one diamond is obviously $\mathcal{A}_m$. Their phases ϕ_m , however, differ: in the one diamond the momentum of the field oscillator is positive, corresponding to motion to the "right" whereas in the other, negative, as appropriate for motion to the "left". We thus arrive at

$$W_m = |\sqrt{\mathcal{A}_m} \exp(i\phi_m) + \sqrt{\mathcal{A}_m} \exp(-i\phi_m)|^2$$

which is identical to equation (7). According to equation (9) the interference-determining phase ϕ_m is fixed by the area in phase space caught between the centre lines of the two states. This phase-fixing area appears as the shaded zone in Fig. 3.

Let us compare and contrast the physics of a squeezed state, as just examined, with the physics of the familiar double-slit experiment. In both cases there are two interfering contributions to the total probability of detecting a specified outcome. In one case the two contributions come from the two slits. In the other case they result from the two distinct areas of overlap in phase space. The phase difference in the double-slit experiment is measured by the difference in optical path length from the centres of the two slits to the point of detection. Likewise, the probability amplitudes for the two contributions to finding m photons in a squeezed state have a phase difference. It is governed by the area in oscillator (x, p) space caught between the initial and final orbits expressed by equation (9). In this sense Young's famous interference experiment is generalized to 'interference in phase space'. The striking effect predicted here still remains to be detected in the realm of quantum optics.

We thank H. Carmichael, R. Y. Chiao, L. Cohen, K. Dodson, M. Hillery, H. J. Kimble, K. Kraus, R. F. O'Connell, M. O. Scully, D. F. Walls and H. Walther for useful and stimulating discussions. In particular, we thank M. Brown for computer evaluation of curves in Fig. 1. Preparation of this article was assisted by the Center for Theoretical Physics, Austin, and by NSF.

Received 20 February; accepted 24 February 1987.

1. Yuen, H. P. *Phys. Rev. A* **13**, 2226–2243 (1976).
2. Yuen, H. P. in *Quantum Optics, Experimental Gravity and Measurement Theory* (eds Meystre, P. & Scully, M. O.) 249–268 (Plenum, New York, 1983).
3. Hollenhorst, J. N. *Phys. Rev. D* **19**, 1669–1679 (1979).
4. Caves, C. M. *Phys. Rev. D* **23**, 1693–1708 (1981).
5. Walls, D. F. *Nature* **306**, 141–146 (1983).
6. Nieto, M. in *Frontiers of Nonequilibrium Statistical Mechanics* (eds Moore, G. & Scully, M. O.) 287–307 (Plenum, New York, 1986).
7. Stoler, D. *Phys. Rev. D* **1**, 3217–3219 (1970); *Phys. Rev. D* **4** 1925–1926 (1971).
8. Slusher, R. E., Hollberg, L. W., Yurke, B., Mertz, J. C. & Valley, J. F. *Phys. Rev. Lett.* **55**, 2409–2412 (1985).
9. Shelby, R. M., Levenson, M. D., Perlmutter, S. H., DeVoe, R. G. & Walls, D. F. *Phys. Rev. Lett.* **57**, 691–694 (1986).
10. Maeder, M. W., Kumar, P. & Shapiro, J. H. in *Proc. of the Joint Meetings: Fourteenth International Conference on Quantum Electronics and Sixth Annual Conference on Lasers and Electro-Optics* San Francisco, California 9–13 June (1986) (in the press).
11. Wu, L.-A., Kimble, H. J., Hall, J. L. & Wu, H. *Phys. Rev. Lett.* **57**, 2520–2523 (1986).
12. Levi, B. G. *Physics Today* **39**, (3), 17–19 (1986).
13. Robinson, A. L. *Science* **233**, 280–281 (1986).
14. Wheeler, J. A. *Lett. math. Phys.* **10**, 201–206 (1985).
15. Schleich, W. & Wheeler, J. A. in *Proc. of the first International Conference on the Physics of Phase Space* (ed. Zachary, W. W.) (Springer, New York, in the press).
16. Bohm, D. *Quantum Theory* (Prentice-Hall, Englewood Cliffs, New York, 1951).
17. Condon, E. U. *Am. J. Phys.* **15**, 365–374 (1947).
18. Herzberg, G. *Molecular Spectra and Molecular Structure, I. Spectra of Diatomic Molecules* 194–204 (van Nostrand, Princeton, 1965).
19. Born, M. in *Struktur der Materie in Einzeldarstellungen* (eds Born, M. & Franck, J.) (Springer, Berlin, 1925).
20. Szegő, G. *Orthogonal Polynomials* (American Mathematical Society, New York, 1939).
21. Sargent, M., Scully, M. O. & Lamb, W. E. *Laser Physics* 242–256; 410–418 (Addison-Wesley, Reading, 1974).
22. Wheeler, J. A. & Zurek, W. H. *Quantum Theory and Measurement* (Princeton University Press, 1983).
23. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. 3 (Addison-Wesley, Reading, 1964).
24. Debye, P. *Physik. Zeitschrift* **28**, 170–174 (1927).
25. Liboff, R. L. *Physics Today* **37**(2), 50–55 (1984).
26. Planck, M. *Ann. Phys.* **50**, 385–418 (1916).
27. Cohen, L. & Zaporovanny, J. *Math. Phys.* **21**, 794–796 (1980).
28. Cohen, L. in *Frontiers of Nonequilibrium Statistical Physics* (eds Moore, G. T. & Scully, M. O.) 97–117 (Plenum, New York, 1986).

A model of solar-cycle effects on palaeoclimate and its implications for the Elatina Formation

J.-Cl. Gérard & L. M. François

Astrophysical Institute, University of Liège, 5, avenue de Cointe, B-4200 Cointe-Liège, Belgium

Recently published analysis of periodicities observed in the characteristics of the annual deposits of a Precambrian (680 Myr ago) lake in South Australia^{1,2} suggest a physical link between ancient solar variability, surface temperature and cyclicity in varve thickness. Solar-cycle effects on the present climate have remained elusive³. The smaller abundance of molecular oxygen and ozone in the Precambrian atmosphere made it possible for solar ultraviolet radiation to interact directly with the troposphere. In those conditions, global climatic perturbations may have been associated with the 11-year solar cycle. We use a radiative–convective–photochemical model to investigate the sensitivity of the global surface temperature to changes in the ultraviolet solar irradiance. We show that the response maximizes for an O₂ atmospheric content of 3×10^{-3} the present atmosphere level. Two alternative hypotheses about the ultraviolet variability are examined. In the first, the solar-cycle variation is restricted to wavelengths below 200 nm. The induced surface temperature fluctuation, ΔT_s , is very small (<0.03 K), except that the amplitude of the ultraviolet modulation is enhanced with respect to average values for solar cycle 21. The second hypothesis assumes that the ultraviolet variability extends up to 300 nm. In this case, a ΔT_s perturbation of several tenths of kelvin is predicted. It is also shown that other local effects such as the amount of water vapour in the troposphere may restrict the conditions in which solar-cycle variations have affected the palaeoclimate.

The clearest evidence of a control of varve thickness by solar-cycle variability has been found in drill-core material from

Table 1 Surface temperature and ozone column variations between solar maximum and minimum conditions for various models at 10^{-3} PAL of O₂ (see text)

Model case	ΔT_s (K)	$\Delta[\text{O}_3]$ (cm ⁻²)	$\Delta[\text{O}_3]/[\text{O}_3]$ (%)
1. Ultraviolet variability for $\lambda < 200$ nm (hypothesis A)	0.02	4×10^{16}	1.8
2. Ultraviolet variability for 30 p.p.m. H ₂ O in the stratosphere	0.02	3×10^{16}	1.4
3. Ultraviolet variability for $\lambda < 300$ nm (hypothesis B)	0.34	5.1×10^{17}	17.6
4. Ultraviolet variability for $\lambda < 200$ nm with flux ratio $r(\lambda)$ multiplied by 20	0.41	5.6×10^{17}	25.5
5. Same as model 4 for a CO ₂ -rich atmosphere ($p_{\text{CO}_2} = 10$ PAL)	0.53	7.2×10^{17}	26.3
6. Same as model 4 with lowered cloud opacity (warm climate)	0.12	1.4×10^{17}	15.5

the Elatina Formation at the Pichi-Richi Pass in the Flinders Ranges (South Australia). The site was presumably the location of a periglacial lake whose melted discharge was strongly controlled by a marked seasonal arid palaeoclimate. The non-random cyclicities found in the 19,000-yr sample and comparison with modern analogues make it possible to interpret the variations of the varve thickness as being due to periodic perturbations of the annual temperature. Time-series analyses of over 1,300 varve sequences have revealed the existence of a 12-yr periodicity in the varve thickness measurements. This period and longer cycles also found in the recent variations of solar activity demonstrate a solar control of the mean annual temperature at the time and location of the Elatina deposits.

One attractive explanation of the larger climatic sensitivity of the ancient troposphere to solar ultraviolet cycles is tied to the low ozone content of the Precambrian atmosphere. Indeed, in the presence of smaller amounts of oxygen present in the atmosphere at the early stage of photosynthetic production, the ozone layer was lower than now. It may therefore be expected that the solar-cycle modulation⁴ observed today in the stratosphere occurred closer to the ground in an oxygen-poor atmosphere. A radiative–convective–photochemical model extending from the ground to 70 km with variable solar ultraviolet input has been developed to assess quantitatively the potential role of this mechanism. The model consistently calculates the vertical distribution of temperature, and major and minor constituents.

The radiative calculation is similar to models previously developed for the present atmosphere⁵. The one-dimensional thermodynamic equation is solved using a forward-time-marching method. The absorption of solar radiation by CO₂, H₂O, O₃ and clouds is calculated using the Lacis and Hansen⁶ formulation, which includes the effects of Rayleigh scattering and surface albedo. The cloud cover is characterized by a single layer with fixed altitude and optical depth.

The variation of surface albedo with the average ground temperature is adopted from the results of the energy-balance model by Wang and Stone⁷. The greenhouse effect due to absorption in the infrared by O₃, CO₂, N₂O and CH₄ is treated following the method of Ramanathan⁸. The H₂O contribution is calculated from the parameterization given by Sasamori⁸. The tropospheric temperature gradient is not allowed to exceed the moist adiabatic lapse rate corresponding to a convective profile. This convective adjustment is made following the method of Manabe and Wetherald⁹. A system of coupled continuity and turbulent flux equations is solved for O_x (O¹D, O, O₃), HO_x (H, OH, HO₂), NO_x (NO, NO₂, NO₃, N₂O₅, HNO₂, HNO₃, N₂O₄), H₂O₂, N₂O, CO and O₂. The following species are assumed to be in photo-

chemical equilibrium at each altitude: H_2CO , CH_3 , CH_3O_2 , H_3CO , CH_3OOH and HCO . The 'family' method is used to avoid the numerical difficulties due to the widely differing chemical time constants of the various constituents. However, at a low mixing ratio of O_2 , the continuity equation is solved separately for H , OH and HO_2 . The individual constituents are partitioned assuming photochemical equilibrium within each family. A total of 90 chemical and photochemical processes are included in the model, and recent compilations of rate coefficients and cross-sections are used¹⁰. Turbulent vertical mixing is parameterized by using a typical profile of the eddy diffusion coefficient in the present atmosphere. Values adopted for ultraviolet solar irradiances are based on recent rocket and satellite measurements. Owing to the feedback between the water vapour and ozone distributions and the thermal structure, numerical integration is alternated between radiative and photochemical calculations until convergence is achieved.

The oxygen level (in the present atmosphere level (PAL)) is fixed by imposing the mixing ratio of O_2 at the ground. For $[\text{O}_2] = 10^{-1}$ PAL and 1 PAL, this mixing ratio is kept constant at each height in the atmosphere, the total pressure being calculated consistently with the temperature profile. For lower oxygen concentrations the continuity equation is solved for O_2 because, owing to chemical processes, it tends to depart from hydrostatic equilibrium. In most cases the water vapour in the troposphere is fixed by assuming that the relative humidity profile follows the parameterization of Manabe and Wetherald. In the stratosphere, the water-vapour mixing ratio is fixed assumed to be that at the tropopause.

Because the Elatina deposits were formed during the Marinoan Glaciation 680 Myr ago^{1,2}, the model cloud layer is fixed at low altitude with an optical thickness $\tau_c = 10$ and a cloud cover $A_c = 0.5$ to produce surface temperatures T_s substantially lower than the present Earth temperature. With these parameters, T_s generally falls in the range 268–275 K.

The ratio $r(\lambda) = F_{\text{max}}(\lambda)/F_{\text{min}}(\lambda)$ between the ultraviolet irradiances at solar maximum and minimum is not well known, even for the modern Sun. This ratio has been measured accurately only for cycle 21. Below 190 nm r increases regularly with decreasing wavelength, except in the vicinity of the Lyman- α line, where r is smaller¹¹. Above 190 nm the published observations are highly divergent¹², owing to difficulties of long-term stability of the instruments. In view of this disparity, two arbitrary wavelength dependences for $r(\lambda)$, covering approximately the whole range of uncertainty, were adopted. The first (hypothesis A) assumed a linear increase from $r=1$ for $\lambda = 200$ nm to $r=2.6$ for $\lambda = 120$ nm, with no variability above 200 nm. In the second (hypothesis B), the ultraviolet modulation extended up to 300 nm, with $r(\lambda)$ being linearly interpolated between $r(120 \text{ nm}) = 2.6$ and $r(300 \text{ nm}) = 1.0$. For both cases, $r = 1.8$ was adopted in the Lyman- α emission. The data for the mean ultraviolet irradiance reported by Brasseur and Simon¹³ were used, together with the value of r to generate $F_{\text{max}}(\lambda)$ and $F_{\text{min}}(\lambda)$. The model was run with each of these irradiances for various oxygen levels. With hypothesis A the surface temperature difference ΔT_s between solar maximum and minimum was always very small, ~ 0.02 or 0.03 K. For example, at $[\text{O}_2] = 10^{-3}$ PAL, $T_s \approx 0.02$ K (Table 1, case 1) and, at $[\text{O}_2] = 10^{-2}$ PAL, $T_s = 0.01$ K. These small values suggest that the first solar-cycle variation of the ultraviolet flux (hypothesis A) is too small to explain the signature of the solar control in the Elatina Formation. To investigate the sensitivity of the calculation to stratospheric water vapour in the ancient atmosphere, a model with 30 p.p.m. H_2O in the stratosphere was run for $[\text{O}_2] = 10^{-3}$ PAL. Table 1 (case 2) shows that the results are not altered by such a change, except that the variation of the ozone column between maximum and minimum is somewhat lower.

With hypothesis B the surface temperature response ΔT_s was much larger. Indeed, as shown in Table 1, for $[\text{O}_2] = 10^{-3}$ PAL, $\Delta T_s = 0.34$, that is, >10 times larger than the corresponding

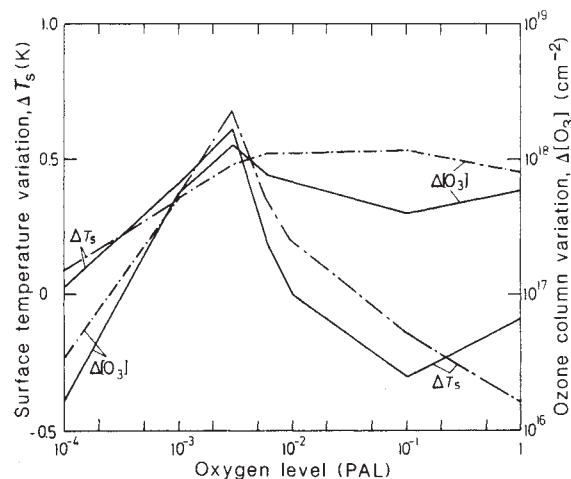


Fig. 1 Variations in surface temperature and total ozone column between solar maximum and minimum as a function of the atmospheric oxygen concentration. Two hypotheses of solar ultraviolet modulation are tested: variability for $\lambda < 300$ nm (dashed-dotted line), and variability for $\lambda < 200$ nm with $r(\lambda)$ enhanced 20-fold (solid line).

value for hypothesis A. Figure 1 shows the dependence of the surface temperature (ΔT_s) and ozone column ($\Delta[\text{O}_3]$) variations between solar maximum and minimum against the oxygen concentration. The temperature response maximizes for $[\text{O}_2] \approx 3 \times 10^{-3}$ PAL, with $\Delta T_s = 0.68$ K. For lower O_2 concentrations ΔT_s shows a regular increase with $[\text{O}_2]$, whereas for higher $[\text{O}_2]$, ΔT_s decreases and even becomes negative. The perturbation in the ozone column increases regularly for $[\text{O}_2] < 3 \times 10^{-3}$ PAL and stabilizes at 10^{18} cm^{-2} for higher $[\text{O}_2]$.

The extent of the solar ultraviolet variability is not well known and it is quite possible that the ratio $r(\lambda)$ has varied from one cycle to another. Thus the possibility that this ratio was greater at the end of the Precambrian era had to be examined. To test the sensitivity of ΔT_s to the magnitude of the ultraviolet fluctuations, the model based on hypothesis A was run for $r(\lambda)$ values increased 20-fold. The results are shown in Fig. 1 as a function of the O_2 concentration. With such an increase of $r(\lambda)$, the magnitudes of ΔT_s and $\Delta[\text{O}_3]$ are similar to those calculated above for hypothesis B. Again the temperature response maximizes for $[\text{O}_2] = 3 \times 10^{-3}$ PAL and changes sign for larger O_2 concentrations.

This dependence on O_2 is very complex and is tied to thermal structure, as well as to the altitude where the ultraviolet energy is deposited. At low $[\text{O}_2]$, the peak of ozone concentration is located at 11 km, independently of the O_2 concentration (in agreement with the model of Kasting and Donahue¹⁴). This altitude corresponds also to the largest $[\text{O}_3]$ fluctuation induced by the solar cycle. For these O_2 concentrations, as Fig. 1 shows, the O_3 column variation $\Delta[\text{O}_3]$ increases with $[\text{O}_2]$. The increase of ΔT_s with $[\text{O}_2]$ is a direct consequence of a greater fluctuation of the O_3 'greenhouse' due to a larger $\Delta[\text{O}_3]$. However, for $[\text{O}_2] > 10^{-3}$ PAL, the ozone peak is progressively shifted to a higher altitude and simultaneously a stratosphere (with a temperature maximum) is developed. The altitude of maximum $\Delta[\text{O}_3]$ (induced by the solar cycle) increases similarly. At $[\text{O}_2] = 1$ PAL, this effect occurs high in the atmosphere, increasing the stratospheric temperature for solar maximum with respect to solar minimum conditions (see Fig. 2). However, such an increase in stratospheric temperature decreases the tropopause temperature and hence the surface temperature, corresponding to a negative ΔT_s in Fig. 1. In other words, an increase of the ozone-mixing ratio in the vicinity of the stratospheric temperature peak generates, in the present model, a decrease of surface temperature, whereas such a perturbation of the ozone profile has the opposite effect lower in the atmosphere (or

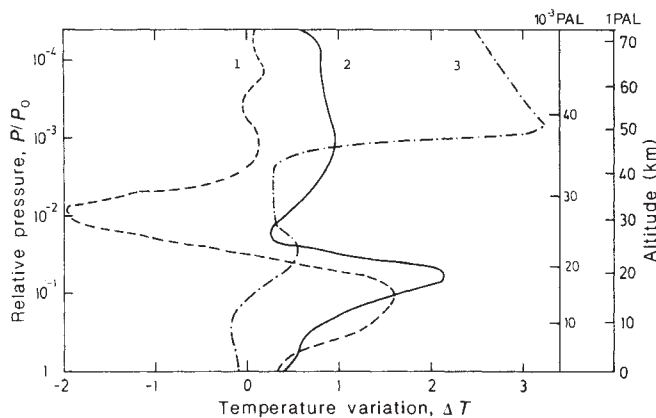


Fig. 2 Vertical profile of atmospheric temperature variation between solar maximum and minimum (P_0 is the surface pressure). 1, Case 3 of Table 1; 2, case 4 of Table 1; 3, 1 PAL, ultraviolet variability restricted to $\lambda < 200$ nm and $r(\lambda)$ enhanced 20-fold. Two approximate altitude scales are given on the right-hand side for 10^{-3} PAL (1, 2) and 1 PAL (3).

similarly when there is no stratosphere). It is not clear whether this result is a general conclusion or whether it is tied to the particular values of some model parameters. It must also be mentioned that the vertical structure of the temperature difference ΔT between maximum and minimum conditions is strongly dependent on the adopted ultraviolet variability (see Fig. 2 for the profile at 10^{-3} PAL).

The carbon dioxide partial pressure was presumably higher in the past than at present¹⁵. Thus, as a sensitivity test, the response of the model to the solar cycle was studied for a CO_2 -rich atmosphere at $[\text{O}_2] = 10^{-3}$ PAL. The ultraviolet variability of model 4 in Table 1 (hypothesis A with $r(\lambda)$ increased 20-fold) has been used. The results indicate that ΔT_s and $[\text{O}_3]$ are somewhat larger than if $p_{\text{CO}_2} = 1$ PAL, but the orders of magnitude are unaffected.

A last case is presented in Table 1, testing the sensitivity of the model to surface temperature. A warm climate corresponds to a wet troposphere. An increase of tropospheric moisture generates, on the one hand, a lower tropospheric ozone content, and, on the other, a less deep penetration of solar ultraviolet. Both effects act to decrease the fluctuation of the ozone column, $\Delta[\text{O}_3]$, during the solar cycle. Consequently, a smaller ΔT_s should be observed. Case 6 in Table 1 corresponds to such a warm and wet climate (~ 291 K), obtained by lowering the cloud opacity in model 4 ($[\text{O}_2] = 10^{-3}$ PAL). The decrease in ΔT_s is pronounced, as a ΔT_s of only 0.12 K is obtained. This result is important because it indicates that the solar-cycle signature in varves would be much more easily detected under cold climate conditions, as during the Elatina deposition.

We do not examine here the possibility that the solar-cycle signal in varves was induced by a simple variation of the solar constant, S . Indeed, precise measurements on the Solar Maximum Mission¹⁶ have shown that the solar constant may vary by several tenths of a per cent over periods of days, owing to the blocking of radiation by sunspots. The same mechanism would imply a mean variation of the solar constant during the 11-year cycle with an amplitude of $\sim 0.1\%$ (ref. 17). In a climate model, the sensitivity to the solar constant is measured by the parameter $\beta = S\Delta T_s/\Delta S$, where ΔT_s is the surface-temperature variation induced by a variation ΔS of the solar constant. With the conditions of the standard model of Table 1, $\beta = 25$ K, which is much smaller than in usual climate models. This low sensitivity is due partly to the negative feedback associated with the use of a moist adiabatic lapse rate in the troposphere¹⁸ and partly to the change in cloud opacity. Thus, with the present model, a 0.1% variation of the solar constant would result in a surface-temperature change $\Delta T_s = 0.025$ K. This figure should be com-

pared with $\Delta T_s \approx 0.02$ K obtained for case 1 in Table 1, that is, for the smallest variability of the ultraviolet irradiance. As both results were derived with the same model, this comparison indicates that the solar ultraviolet variation during a solar cycle is at least as important as the solar constant variation to explain the surface-temperature changes.

A key question is whether the ΔT_s values calculated in this study are large enough to be observable through varve thickness measurements. This is probably true for the largest ΔT_s shown in Table 1, because a given variation of the global Earth surface temperature would generate local variations, at high latitudes, more than five times larger, as suggested by models of both modern and Cretaceous climates^{19,20}. By contrast, if it is assumed that the ultraviolet variability has a modern amplitude and is restricted to $\lambda < 200$ nm, the calculated ΔT_s is very small for all O_2 levels and its ability to produce a modulation of varve thickness remains questionable. However, note that, first, the low β in the present study possibly indicates a low sensitivity of the radiative model to external perturbations (such as variations of the O_3 column). Actually, this sensitivity could have been much larger, because both the cloud feedback and the surface albedo response to mean global temperature are not well known. Second, Sonett and Williams²¹ have found a positive correlation between the sunspot index and varve thickness in recent glacial lake deposits. This observation proves that very small ΔT_s are large enough to produce at least a small signal in varves. This positive correlation implies that the surface temperature is currently higher at solar maximum. This result is puzzling because under present conditions, both a variation of the solar constant and of the ultraviolet irradiance would result in a negative correlation.

Nevertheless, whatever the actual explanation to this problem may be, our calculations show that the solar ultraviolet variability may have played a major role in modulating surface climate. Indeed, the temperature response ΔT_s strongly depends on the O_2 level and maximizes for $[\text{O}_2] = 3 \times 10^{-3}$ PAL. Several authors^{22,23} suggest an atmospheric O_2 concentration of several per cent of PAL at the time of the Elatina deposition. For such O_2 abundances, the present model yields slightly negative ΔT_s values with magnitudes comparable with the results at $[\text{O}_2] = 1$ PAL. Consequently, if the few per cent O_2 concentration is accepted, it seems difficult to invoke the small O_2 amount of the Precambrian atmosphere to explain the specificity of the Elatina Formation. These deposits contain, indeed, the most obvious signature of solar cyclicity: in modern sediments, the solar signal is very weak and even seems absent in some varve series of the Eocene Green River Formation²⁴.

A conclusion about the specificity of the Elatina Formation must await further studies on this topic. The main contribution of this work is to demonstrate that the climatic response of solar cyclicity is highly dependent on the atmospheric O_2 concentration and may be enhanced for cold climates.

J.C.G. is supported by the Belgian National Foundation for Scientific Research (FNRS). We thank R. G. Roble for valuable discussions.

Received 23 October; accepted 31 December 1986.

- Williams, G. E. & Sonett, C. P. *Nature* **318**, 523–527 (1985).
- Williams, G. E. *Scient. Am.* **225**, 88–96 (1986).
- Pitcock, A. B. *Q. J. R. met. Soc.* **109**, 23–55 (1983).
- Chandra, S. *J. geophys. Res.* **89**, 1373–1379 (1984).
- Ramanathan, V. *J. atmos. Sci.* **33**, 1330–1346 (1976).
- Lacis, A. A. & Hansen, J. E. *J. atmos. Sci.* **31**, 118–133 (1974).
- Wang, W. C. & Stone, P. H. *J. atmos. Sci.* **37**, 545–552 (1980).
- Sasamori, T. *J. appl. Met.* **7**, 721–729 (1968).
- Manabe, S. & Wetherald, R. T. *J. atmos. Sci.* **39**, 1189–1205 (1967).
- JPL Publication 85–37 (Pasadena, California, 1985).
- Brasseur, G. & Solomon, S. *Aeronomy of the Middle Atmosphere* (Reidel, Dordrecht, 1984).
- Garcia, R. R., Solomon, S. & Roble, R. G. *Planet. Space Sci.* **32**, 411–423 (1984).
- Brasseur, G. & Simon, P. C. *J. geophys. Res.* **86**, 7343–7362 (1981).
- Kasting, J. F. & Donahue, T. M. *J. geophys. Res.* **85**, 3255–3263 (1980).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Ocean* (Princeton University Press, 1984).
- Willson, R. C., Gulkis, S., Janssen, M., Hudson, H. S. & Chapman, G. A. *Science* **211**, 700–702 (1981).

17. Hoyt, D. V. & Eddy, J. A. *NCAR tech. Note* TN-194+STR (1982).
18. Lindzen, R. S., Hou, A. Y. & Farrell, B. F. *J. Atmos. Sci.* **39**, 1189-1205 (1982).
19. Wang, W. C., Wuebbles, D. J., Washington, W. M., Issacs, R. G. & Molnar, G. *Rev. Geophys.* **24**, 110-140 (1986).
20. Barron, E. J. & Washington, W. M. *J. geophys. Res.* **89**, 1267-1279 (1984).
21. Sonett, C. P. & Williams, G. E. *J. geophys. Res.* **90**, 12019-12026 (1985).
22. Walker, J. C. G. *Evolution of the Atmosphere* (Macmillan, New York, 1977).
23. Cloud, P. *Mem. geol. Soc. Am.* **161**, 245-251 (1983).
24. Crowley, K. D., Duchon, C. E. & Rhi, J. *J. geophys. Res.* **91**, 8637-8647 (1986).

Synthesis of fibres by growth-packing

Han Sik Yoon

Fiber and Polymer Synthesis Laboratory, Korea Advanced
Institute of Science and Technology, PO Box 131,
Dong Dae Moon, Seoul, Korea

Molecular growth and self-ordering, the process that gives rise to wool, cotton and other fibrous biological materials, has until now been observed only in nature. Controversy exists as to whether these fibrillar elements have fundamental significance in relation to bio-synthesis or whether they are simply secondary aggregation phenomena after molecular bonding has been completed¹. Since the initial attempt by Staudinger in 1927², numerous attempts at simulation of such growth-packing of artificial fibres have been made without success; Pennings *et al.*³ effected chain packings with polyethylene in a xylene solution, but without molecular growth. Here I report the first synthesis of fibres from poly-*p*-phenylene-terephthalamide (PPT-A) by means of a growth-packing mechanism⁴ (Fig. 1). They were grown in a stationary gel phase, containing active oligo-PPT-A, heterocyclic tertiary amines, alkali metal cations and a polar solvent. Molecular growth seemed to be attained in an orderly gel phase, in which the critical chain distance between polymer molecules was maintained at 2.4 nm on average by solvent bridges, built by induced polarizing forces between the polar groups in the chains. After the growth of the molecules, the chain packing took place, with lateral movement of the grown molecules due to the collapse of the solvent bridges. The fibres thus formed have some similarity to native flax or ramie fibres.

PPT-A growth-packing experiments were done as follows: 12.24 g of terephthaloylchloride (TPC) was added to a vigorously stirred solution of 6.48 g of *p*-phenylenediamine (PPD) (15 ml of pyridine and 6.0 g of LiCl was dissolved in 240 ml of di-methylacetamide (DMAc) at 25 °C). The almost instantaneous polymerization, accompanied by a sharp increase of viscosity, turned the reaction mixture into a hard, agar-like gel within 10 s. When this was allowed to stand still for several hours, fibrils formed in the gel phase with partial release of DMAc. The molecular mass of PPT-A increased to 380% of that of the oligo-PPT-A and the parallel orientation of the oligo-chains from the stirring was maintained. The molecular weight did not increase further once the fibril formation began within the gel phase. After removing DMAc, the entire reaction mixture became a mass of fibrils and, when sliced into fine strands, PPT-A fibres having secondary structures similar to native cotton or linen were obtained (Fig. 2a). If the reaction conditions were varied slightly, the gel did not form and the reaction terminated by precipitating PPT-A from DMAc at an early stage.

The average molecular weight of the PPT-A measured⁵ at the onset of gelation was ~15,000, whereas after the completion of the reaction it reached ~57,000 with a narrow, molecular weight distribution. This is unusual for a polymerization reaction.

Wide-angle X-ray diffractograms of the material show typical fibre reflection patterns (Fig. 2c). An orientation angle of 17-22° was measured from the diffraction data. The cross-section of PPT-A fibrils is characteristically elliptical, and its dark-field

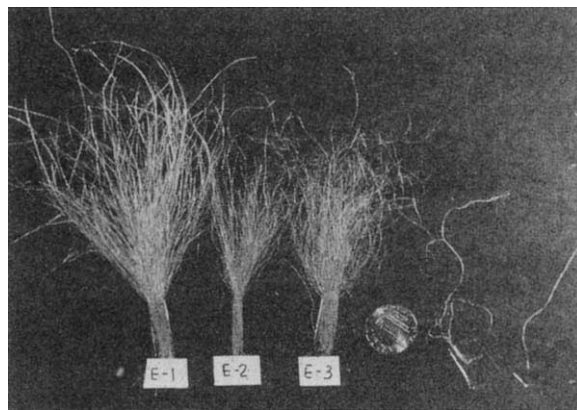


Fig. 1 Growth-packed poly-*p*-phenyleneterephthalamide (PPT-A) fibres with secondary structures similar to native flax or ramie fibres; these fibres can be split again to finer ones.

transverse electron microscope (TEM) image (Fig. 2b) shows a banded structure.

Single-crystal, X-ray diffractometry of 4,4'-diaminophenyl-terephthalamide (Fig. 3a), a fragment of the PPT-A molecule, gave more precise information on the molecular conformation of the PPT-A than was hitherto known^{6,7}; —CO— and —NH— groups in the molecular chains run in opposite directions and the plane made by these groups does not bisect the angle between the two phenyl rings (Fig. 3b). This gives the polarity of each segment as the twist may be either clockwise or anti-clockwise. It should be possible to construct a rigid and fairly flat PTA-A molecule by joining segments of different polarities.

We view the gel-phase structure as follows: the oligo-PPT-A chains, having about 60 repeating units of the above molecular conformation, are arranged parallel to the mechanical shearing direction and are aligned end-to-end by electrostatic interactions. There is a relative chain shift of 1/2 of the repeating unit and —CO— and —NH— groups in the oligo-PPT-As are interconnected laterally with a large number of DMAc bridges of definite length. These bridges are constructed by molecular alignment of DMAc due to induced polarizing forces from the polar groups, acting together with lone-pair electron compounds and alkali-metal cations. The structural model of the gel phase emerging from this consideration is a rectangular three-dimensional lattice structure, and subsequent chemical bonding of the oligo-PPT-A chains is obtained by dehydrochlorination, that speeds up because of the favourable change of activation energy level of the oligo-chains.

The completion of the molecular growth leads to the collapse of the DMAc bridges from the centre by steps and this can propagate instantly throughout the gel phase. Thus DMAc molecules are expelled, leaving many fibrils and microcleavages between them.

The average intermolecular distance of the PPT-A is calculated to be ~2.4 nm, based on the experimental results, which include the polymer/solvent ratio, gel-phase density, and volume shrinkage of the gel-fibril after the solvent removal. The interchain dimension that provides the necessary space for the molecular growth/packing is maintained by solvent bridges built by induced polarizing forces. The number of DMAc molecules involved in the formation of a solvent bridge was calculated to be 22, because the unit cell of the gel phase (2.4 nm × 2.4 nm × PPT-A repeating unit) contains 44 DMAc molecules and 2 DMAc bridges.

The thickness and the molecular mass of the PPT-A fibril is dependent on the temperature of the gel phase before the fibril formation: the higher the temperature the finer the fibrils and the lower the molecular weight. This result is quite natural because the solvent bridges are temperature-sensitive.

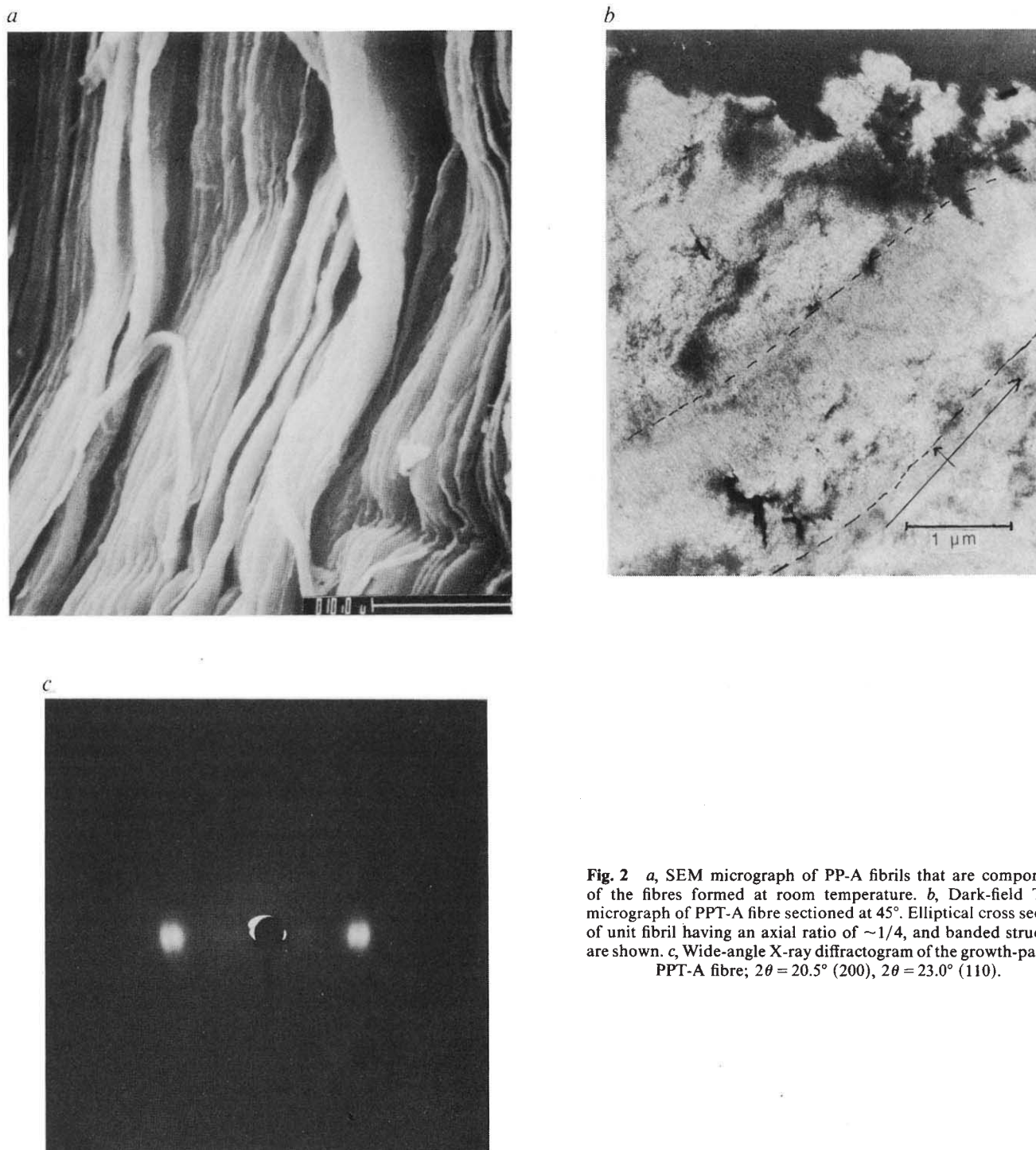


Fig. 2 *a*, SEM micrograph of PP-A fibrils that are components of the fibres formed at room temperature. *b*, Dark-field TEM micrograph of PPT-A fibre sectioned at 45°. Elliptical cross section of unit fibril having an axial ratio of $\sim 1/4$, and banded structure are shown. *c*, Wide-angle X-ray diffractogram of the growth-packed PPT-A fibre; $2\theta = 20.5^\circ$ (200), $2\theta = 23.0^\circ$ (110).

As the DMAc bridges collapse, the PPT-A chains are brought closer together and ultimately chain packing takes place. The self-ordering process of the polymer characterized by this lateral chain movement during packing creates a unique, structural patterns of the PPT-A fibre. First, the numerous fibrils contained in the PPT-A fibre form one continuous three-dimensional network as is obvious from the mechanism of their formation. Second, there exists an intimate, structural relation between each of the fibrils; the $-\text{CONH}-$ and phenyl rings arrange themselves laterally in an ordered manner across the microcapillaries throughout the fibrils. The $-\text{CONH}-$ groups are not occluded but exposed toward the outside of the fibril from the amorphous regions through the microcleavages (except those

in crystalline regions). Therefore the $-\text{CONH}-$ groups are vulnerable to attack by chemicals from outside. The proof is the transverse, straight line of the $-\text{CONH}-$ splitting edge, found in alkali-hydrolyzed PPT-A fibres, that is observed by scanning electron microscopy (SEM) (Fig. 4). Third, because the fibrils are interlocked by the crystalline regions, no independent unit fibril can exist. Fourth, there exists clear bidirectionality in the chain packing: one direction is that of the packing of polar groups due to hydrogen bonding, and the other is that of the packing of non-polar groups due to the van der Waals force. In the first direction, there is mainly compact packing, whereas along the other, there is compact or loose packing depending on the orientation of the groups or the segmental polarity of

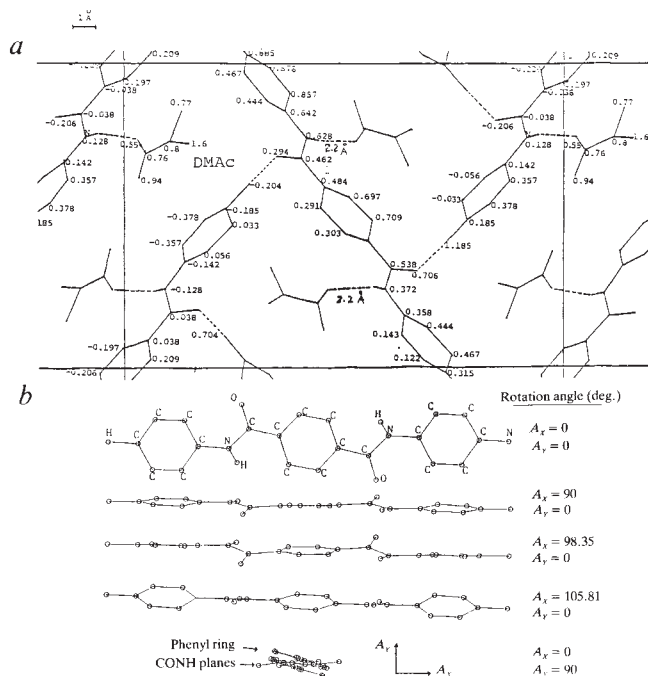


Fig. 3 *a*, 4,4'-Diaminophenylterephthalate (DPT) unit cell obtained by X-ray diffractometry. The single crystal is formed by dimethylacetamide (DMAc) and —NH— groups in the DPT molecule located 0.22 nm from the oxygen atoms of DMAc. The angle between the terephthalic and diamine phenyl ring is only 8.35° and the planes of —CO— and NH that run in opposite directions make an angle of 7.46° with the diamine phenyl ring. *b*, Molecular features of DPT obtained by computer graphics with spatial X, Y, Z data upon rotation of various angles and directions.

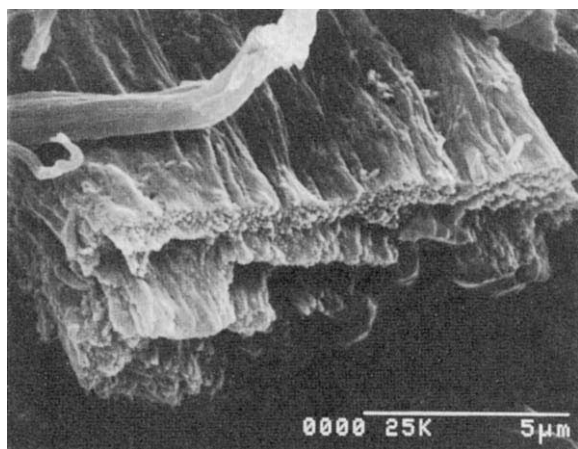


Fig. 4 Transverse straight line of splitting edges made by a PPT-A fibre when it is hydrolysed in 20% KOH aqueous solution. This seems to indicate that all of the —CONH— groups in the fibrils are arranged parallel to the transverse direction.

the groups. These two types of packing alternate along the chain axis. It seems quite natural to speculate that the coincidence of compact packing made by these two forces only results in crystalline regions, whereas the heterocombination of compact and loose packing results in semicrystalline regions. Combination of two loose packings results in amorphous regions. Sheet-like crystalline and amorphous regions seem to be formed by this mechanism. Fifth, the cross-section of the fibril tends to be elliptical in form because of the difference in the shrinking ratio of the fibril radii depending on the packing directions of the chains. Consequently, the polar groups pack along the minor

axis of the cross-section, whereas the non-polar groups pack along the major axis. Finally, the eventual shape of the fibre is a helically twisted form because the molecular aggregation forces operating over very short distances tend to hold the chains tightly together. This is supported by an experiment in which solvent-drained PPT-A gel strands always make a helically twisted fibre.

The chain-packing mechanism found in formation of PPT-A fibrils may be a universal one that applies to formation of all fibrous materials in nature because there are polar groups such as —OH— and —O— in cellulosic fibrils, and —CONH— groups in protein fibrils equivalent to that of the PPT-As. Of course the polar solvent that makes interbridges between the polar segments is H_2O .

I thank Dr D. J. Johnson of the University of Leeds for his editorial suggestions, Mr R. Hagege of the Institut Textile de France for the TEM micrograph, Dr H. C. Shin of Seoul National University for the single X-ray crystallography, and the research staff in my laboratory.

Received 23 July 1986; accepted 2 January 1987.

1. Burgess, J. *Nature* **278**, 211 (1979).
2. Staudinger, H., Jöhner, H., Singer, R., Mie, G. & Hengstenberg, J. *Z. phys. Chem.* **126**, 425–448 (1927).
3. Pennings, A. J. Kiel, A. M. *Kolloid Zh.* **205**, 160–163 (1965).
4. Yoon, H. S. & Son, T. W. *J. high perf. Text.* Vol. 4, Number 12, 1–3 (1984).
5. Arpin, M. & Slragiella, C. *Polymer* **18**, 591–598 (1977).
6. Notholt, M. G. & Van Aartsen, J. J. *J. Polym. Sci., Polymer Letters edn* **11**, 333–337 (1973).
7. Northolt, M. G. *Eur. Polymer J.* **10**, 799–804 (1974).

Dynamic ordering of aggregated mesomorphic macromolecules

Ivan Cabrera & Valeri Krongauz

Department of Structural Chemistry,
The Weizmann Institute of Science, Rehovot, 76100 Israel

The guest–host interactions of pleochroic dyes dissolved in liquid crystals are frequently used to obtain information about the structure (usually order parameter) of the mesophases¹. Almost no attention has been paid, however, to the possible influence of dipolar or highly polarizable dye molecules on the liquid crystal solvent, mainly because the solubility of the dyes is very low and no significant macroscopic effect is expected in dilute solution (usually 0.1–1%). The synthesis of liquid crystal copolymers containing dye and mesogenic units² opens a way to incorporate a dye in amounts large enough to change the structure of the mesomorphic system. Here we report structural transformations occurring in side chain liquid crystal copolymers containing thermochromic spiropyran side groups. The spiropyran to merocyanine dye conversion occurs on heating of the copolymers. Physical crosslinking of the macromolecules due to aggregation of the dye moieties gives rise to a network formation, responsible for the appearance of a new rheo-optical effect, observed above the clearing point.

Earlier, we reported^{3,4} a new type of crystallization of atactic vinyl polymers bearing spiropyran side groups, by swelling in polar solvents. Solvatochromic spiropyran–merocyanine conversion and self-assembly of merocyanine groups into molecular stacks are the driving forces for this crystallization. Figure 1 shows spiropyran–merocyanine conversion in a polymer.

We believe that the high degree of crystallinity of the polymers (up to 40%) was achieved because the cooperative spiropyran–merocyanine conversion occurs step-by-step along the polymer chains, and we call this process “zipper crystallization”. The tendency of the merocyanine groups to self-assembly is so strong that even in 1:1 spiropyran methacrylate–methyl methacrylate copolymer it proceeds rather fast, though it does not result in crystalline order⁵.

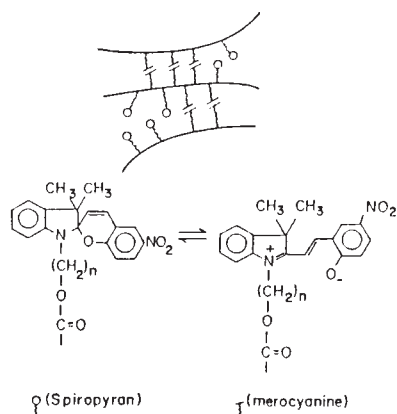


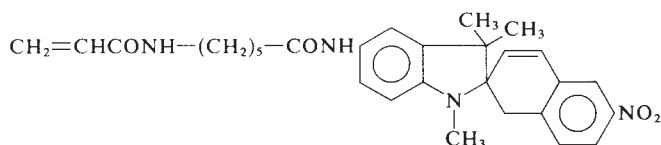
Fig. 1 Spiropyran-merocyanine conversion and aggregation.

The assumed structure of the assemblies resembles the structure of side chain liquid crystal polymers, though the factors that give rise to these structures are completely different: the merocyanine stacks, formed in the course of swelling, belong to the so-called H-type, which is characterized by antiparallel dipole-dipole interactions of the molecules⁶. The liquid crystalline structure on the other hand is determined mainly by the steric factor, connected with the geometrical anisotropy of the mesogenic groups⁷.

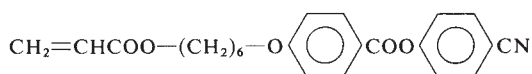
One can expect that in copolymers containing both spiropyran and mesogenic side groups the competition between merocyanine interactions and liquid crystalline ordering of mesogenic groups will give rise to new interesting structural changes of the mesophase, taking into account the fact that spiropyran can be effectively converted to merocyanine also by heat.

Recently, we have shown that the combination of thermochromic spiropyran and mesogenic groups in a low molecular weight molecule gives a new kind of mesophase, quasi-liquid crystals^{8,9}.

In order to make spiropyran groups more compatible with mesogenic ones a new spiropyran acrylate



monomer was synthesized, in which the acrylic group is attached through a flexible spacer to the indoline part of the spiropyran. Copolymerization with an acrylic monomer containing a mesogenic group¹⁰



gives a liquid crystal copolymer with thermochromic properties, which turns red on heating. The homopolymer of the mesogenic acrylate gives a nematic liquid crystal structure¹⁰. The synthesis of the spiropyran monomers and preparation of copolymers will be described elsewhere. The average molecular weight of the polymers, estimated by gel permeation chromatography was 2.5×10^4 for the liquid crystal homopolymer and 8×10^3 for the copolymer containing 22% of spiropyran units.

Observation with a polarizing microscope revealed that the clearing points of the copolymers are lower the higher the content of spiropyran units in the copolymer, going down from 129 °C (homopolymer) to 112 °C and 86 °C (10 and 22 molar % of spiropyran, respectively). The copolymer with 46 molar %

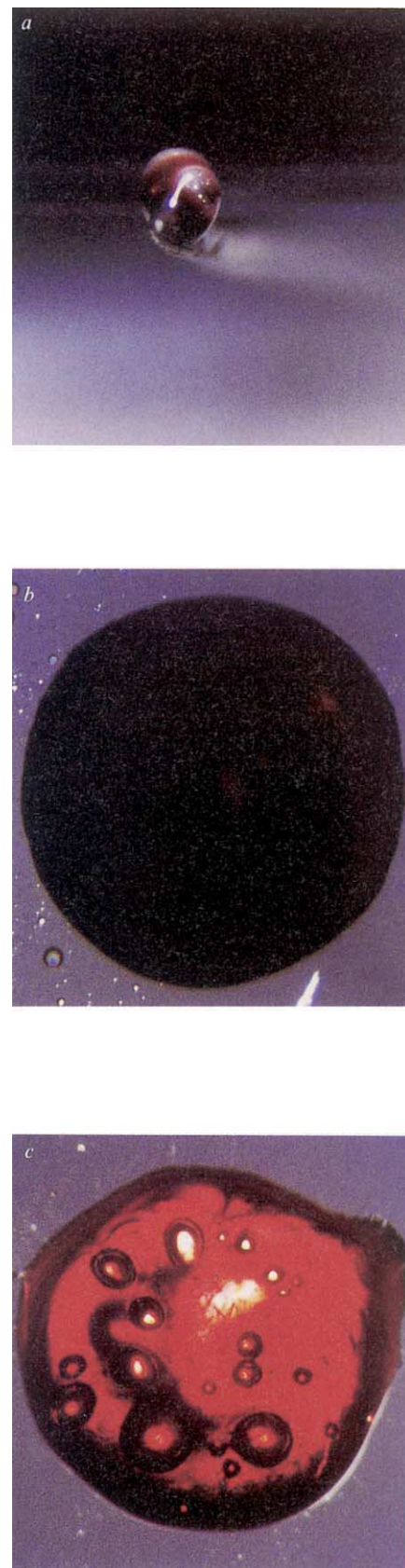


Fig. 2 'Transient birefringence' in droplets of 22 molar % copolymer at 95 °C. *a*, A droplet observed without crossed polarizers ($\times 15$). *b*, Undisturbed droplet observed with crossed polarizers ($\times 60$). *c*, The disturbed droplet between crossed polarizers ($\times 60$).

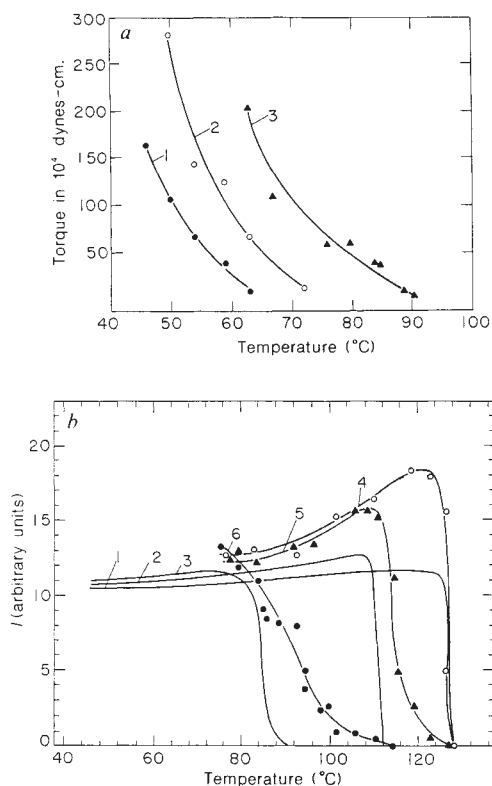


Fig. 3 *a*, The dependence of the torque on temperature: 1, homopolymer; 2, 10 molar % copolymer; 3, 22 molar % copolymer. *b*, Transmitted polarized light intensities (I) against temperature. 'Static' regime: 1, homopolymer; 2, 10 molar % copolymer; 3, 22 molar % copolymer; 4-6, as 1-3, but under 'dynamic' regime.

of spiropyran does not give a mesophase. The most remarkable feature of the films formed by the copolymers with low concentration of spiropyran above the clearing point is their very strong transient translucence between cross polarizers when they are squeezed between two glass slides or even lightly touched with the tip of a spatula.

The liquid crystal homopolymer, which does not contain spiropyran, does not exhibit this effect.

Usually such instant birefringence during mechanical disturbance is considered as an indication of homeotropic (orthogonal to solid surface) orientation of mesogenic molecules¹¹. Therefore, one might conclude that what we observe as a clearing point is not a nematic-isotropic transition but a re-orientation of mesogenic groups of macromolecules from parallel to perpendicular to the surface.

To check this we treated the glass surface with cremophor, nylon and 1-dodecanol^{12,13} which promote respectively the planar and homeotropic alignment of liquid crystals. We found no effect of the surface on the 'transient translucence'. A decisive experiment was performed with droplets of the copolymers in the isotropic phase having a relatively small area connected with a surface (see Fig. 2). For these droplets again even a very gentle touch gave remarkable flash of birefringence though the effect of the surface must be insignificant. This suggests that the transient brightening is caused by at least partial restoration of liquid crystalline order induced by mechanical disturbance, and not by disturbance of existing homeotropic liquid crystalline order.

A parallel-disk type rheometer with transparent glass disks (diameter 1 cm) was built for measurements of the rheo-optical properties of a film as a function of temperature. The measurements were conducted at constant thickness of the polymer films (12 μm) and shear rate (290 s^{-1} at the rim of the device). The

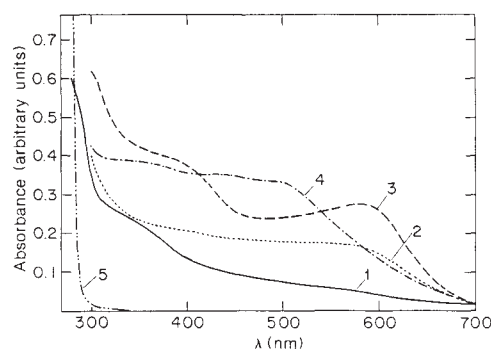


Fig. 4 1-4, Absorption spectra of 22 molar % copolymer films: 1, amorphous film, room temperature (26 $^{\circ}\text{C}$); 2, 82 $^{\circ}\text{C}$; 3, 95 $^{\circ}\text{C}$; 4, swelled film at room temperature; 5, absorption spectrum of homopolymer at room temperature.

torque required to rotate the upper disk was measured as a function of the polymer melt temperature (Fig. 3*a*). The torques of both the liquid crystalline homopolymer and 10 and 22 molar % copolymer decline sharply above the glass transition points (33, 39 and 50 $^{\circ}\text{C}$, respectively). In this region the flow character of the fluid is non-newtonian and very complex. Therefore, the data were not amenable for conversion to viscosity units¹⁴. However, since the viscosity is a monotonically increasing function of the torque¹⁴ we can conclude that the viscosity of the copolymer is much higher than that of the homopolymer at the same temperature. A drastic increase of viscosity upon the introduction of only 10 molar % of the thermochromic groups indicates a very substantial effect of these groups on the interaction between macromolecules. Earlier we showed¹⁵ that spiropyran-merocyanine conversion in melt of low molecular weight spiropyran containing mesogenic groups brings about contraction of the fluid, due to aggregation of the merocyanine formed in the melt. One may assume that in the case of the copolymer the interaction of the merocyanine groups, formed on heating, gives rise to the aggregation of macromolecules in a network with high viscosity.

The polarized light intensity-temperature relationships for the liquid crystalline homopolymer and 10 and 22 molar % copolymers with spiropyran are shown in Fig. 3*b*. The clearing points are indicated by a sharp fall of 'static' birefringence. In the dynamic regime the transmitted light intensity reaches a maximum near the clearing points. This can be explained by a decrease of the number of defects with viscosity decrease and domain growth induced by shear¹⁶. At low temperatures the viscosity becomes so high that it did not allow us to perform the shearing measurements. The viscosity increases also with the spiropyran content: 46 molar % copolymer forms a red, tar-like material on melting, which gives neither mesophase nor dynamic birefringence.

The dynamic birefringence of the homopolymer disappears at the clearing point almost as sharply as does the static one, while that of the copolymers extends far beyond the clearing points, though it declines after that point gradually. The range between the clearing point and the temperature where the dynamic birefringence is not observed anymore increases with the spiropyran content in the copolymers. In other words, even though the spiropyran-merocyanine groups disturb the mesophase thermal stability, they promote the dynamic restoration of an ephemeral order in the isotropic phase.

The differential scanning calorimetry measurements give endothermic peaks at temperatures that coincide with the microscopic observation of the clearing points. This confirms that the static birefringence disappearance relates to the mesophase-isotropic phase transition.

The electronic absorption spectra of the copolymers (see Fig. 4) show an increase of visible absorption with temperature

rise. The transition from amorphous to liquid crystalline phase is accompanied by formation of a broad plateau in the range 460–585 nm. The optical density in this range does not change with temperature up to the clearing point. The transition from mesophase to isotropic phase coincides with the appearance of an absorption band with a $\lambda(\text{max}) = 585$ nm (a correction for the light scattered by the mesophase was made when the spectra were plotted). The band vanishes reversibly on cooling.

Comparison of the spectra with the spectral data obtained by us earlier^{3,8,17} brings us to the conclusion that the broad absorption in the range 460–585 nm is connected with association of merocyanine molecules in molecular stacks (probably very short ones, mostly dimers). The absorption of these associates overlaps apparently with the absorption of isolated spiropyran ($\lambda(\text{max}) = 370$ nm) and merocyanine ($\lambda(\text{max}) = 585$ nm) side groups. This confirms our theory that the elevated viscosity of the copolymer melt may be explained by such association which results in formation of a network. The reversible absorption band with $\lambda(\text{max})$ at 585 nm, characteristic of non-aggregated merocyanine groups, becomes pronounced above the clearing point when formation of the network is completed and restricted segmental mobility hinders further merocyanine association.

In order to stimulate the merocyanine association at room temperature we swelled 22 molar % copolymer film in chloroform. The film was previously heated in order to form a birefringent texture. After swelling and drying, the static birefringence disappeared but a distinct dynamic birefringence was observed even at room temperature. The film acquired a high viscosity and a deep red colour with absorption plateau from 380 to 520 nm (Fig. 4). At longer wavelengths the absorption goes down, which indicates that the contribution of isolated merocyanine moieties to the absorption is negligible. This speaks in favour of merocyanine association as a factor leading to dynamic birefringence.

To explain the results let us consider the effect of the liquid crystal-isotropic phase transition on the main chain conformation. According to Kuzma¹⁸, in the liquid crystal phase the parallel arrangement of mesogenic groups affects the conformation of the macromolecular backbone. Above the clearing point this effect of the mesogenic side groups becomes insignificant and the macromolecules accept a conformation, which provides maximum entropy, and is in general different from the conformation in the mesophase. Formation of the transient liquid crystalline order by shearing or other mechanical perturbation results not only in restoration of parallel arrangement of the mesogenic groups but in a certain reconstruction of the main chain conformation, decreasing entropy.

Presumably, formation of the network favours the preservation by macromolecules of the conformation acquired in the mesophase even above the clearing point due to a more rigid structure. This makes the dynamic ordering easier.

Earlier, Finkelmann *et al.*¹⁹ observed the occurrence of birefringence above the clearing point on stretching side chain liquid crystal polysiloxane elastomers. This observation is consistent with our conclusion that dynamic birefringence stems from the macromolecular aggregation.

We thank Professor M. Cohen for his constant interest in this work and valuable comments on the manuscript. Support from the Minerva Foundation, Munich, FRG, the US-Israel Binational Science Foundation and the US Army European Research Office (contract no. DAJA45-C-0030) is gratefully acknowledged.

Received 23 September 1986; accepted 28 January 1987.

1. Cox, R. J. *Molec. Cryst. Liq. Cryst.* **55**, 1–32 (1979).
2. Ringsdorf, H., Schmidt, H.-W., Baur, G. & Kiefer, R. in *Recent Advances in Liquid Crystalline Polymers* (ed. Chapoy, L.) 253–260 (Elsevier, London, 1985).
3. Krongauz, V. A. & Goldburt, E. S. *Macromolecules* **14**, 1382–1386 (1981).
4. Goldburt, E. S., Shvartsman, F. P. & Krongauz, V. A. *Macromolecules* **17**, 1876–1878 (1984).
5. Wismontski-Knittel, T. & Krongauz, V. A. *Macromolecules* **18**, 2124–2126 (1985).
6. Sturmer, D. M. & Heseltine, D. W. in *The Theory of the Photographic Process* (ed. Jamies, H.) Chs 7, 8 (Macmillan, New York, 1977).

7. Flory, P. J. in *Polymer Liquid Crystals* (eds Ciferri, A., Krigbaum, W. & Meyer, R.) Ch. 4 (Academic, New York, 1982).
8. Shvartsman, F. P. & Krongauz, V. A. *Nature* **309**, 608–611 (1984).
9. Shvartsman, F. P., Cabrera, I. R., Weis, A., Wachtel, E. & Krongauz, V. A. *J. phys. Chem.* **89**, 3941–3946 (1985).
10. Portugall, M., Ringsdorf, H. & Zentel, R. *Makromolek. Chem.* **183**, 2311–2321 (1982).
11. Kelker, H. & Hatz, R. *Handbook of Liquid Crystals* Ch. 1 (Chemie, Berlin, 1980).
12. De Jeu, W. H. *Physical Properties of Liquid Crystalline Materials* 23 (Gordon and Breach, New York, 1980).
13. Patel, J. S., Leslie, T. M. & Goodby, J. W. *Ferroelectrics* **59**, 137–144 (1984).
14. Bird, R. B., Armstrong, R. C. & Hassager, O. *Dynamics of Polymeric Liquids* Vol. 1, Ch. 4 (Wiley, Inc. New York, 1977).
15. Cabrera, I., Shvartsman, F., Veinberg, O. & Krongauz, V. *Science* **226**, 341–343 (1984).
16. Asada, T., Muramatsu, H., Watanabe, R. & Onogi, S. *Macromolecules* **13**, 867–871 (1980).
17. Eckhardt, H., Bose, A. & Krongauz, V. *Polymer* (submitted).
18. Kuzma, M. R. *Molec. Cryst. Liq. Cryst.* **101**, 351–366 (1983).
19. Finkelmann, H., Kock, H., Gleim, W. & Rehage, G. *Makromolek. Chem. rapid Commun.* **5**, 287–293 (1984).

Obsidian provenance determination by back-scattered electron imaging

James H. Burton & David H. Krinsley

Department of Geology, Arizona State University, Tempe, Arizona 85287, USA

Obsidian glass is a significant component of the lithic artefact assemblage of primitive cultures; the ability to characterize and correlate artefacts with specific sources provides significant information about prehistoric patterns of procurement and exchange¹. Because visual and light-optical microscopic methods have proved ineffective in characterizing obsidian, chemical methods such as X-ray fluorescence² (XRF) and instrumental neutron-activation analysis³ (INAA) are usually used to try to determine provenance. Here we show that electron microscopy, specifically back-scattered electron (BSE) imaging, may be used to characterize obsidian petrographically. The variety and number of BSE-observable features are generally sufficient to petrographically 'fingerprint' individual obsidian sources. Sources are also sufficiently homogeneous so that BSE characterization can determine provenance for individual artefacts of archaeological obsidian.

BSE microscopy uses reflected (typically 10–30 keV) high-energy electrons, the number of which is proportional to the mean atomic number of the material under the incident electron beam. This dependence on mean atomic number generates compositionally dependent images. A phase of relatively high mean atomic number, such as iron oxide, reflects more electrons, and thus appears brighter in BSE images, than a phase of relatively low mean atomic number, such as quartz. This compositional imaging capability of BSE permits one to study on micrometre to centimetre scales the chemical and textural relationships among and within minerals and other inorganic materials^{4–6}.

Although the BSE image depends on composition and shows compositional variations, it does not indicate the specific elemental components, so an energy-dispersive X-ray analyser (EDX) is used simultaneously with BSE analysis. When EDX data are combined with the BSE image, individual grains can be chemically identified and variations of chemical composition within individual mineral grains can be examined.

The BSE technique is also rapid and virtually non-destructive. No damage is done to the specimen during the BSE examination and sample preparation is minimal. Because BSE imaging is a technique for surface analysis, petrographic thin sections are not required. Although the technique requires a flat surface, obsidians can be analysed by polishing a facet of less than 1 cm² on the sample and placing the entire piece in the electron microscope. Also, although source characterization requires considerable effort and time, once a database exists against which artefact data can be compared, the BSE examination and source assignment can be made in a matter of minutes. Thus the technique is comparable in efficiency to rapid trace-element methods.

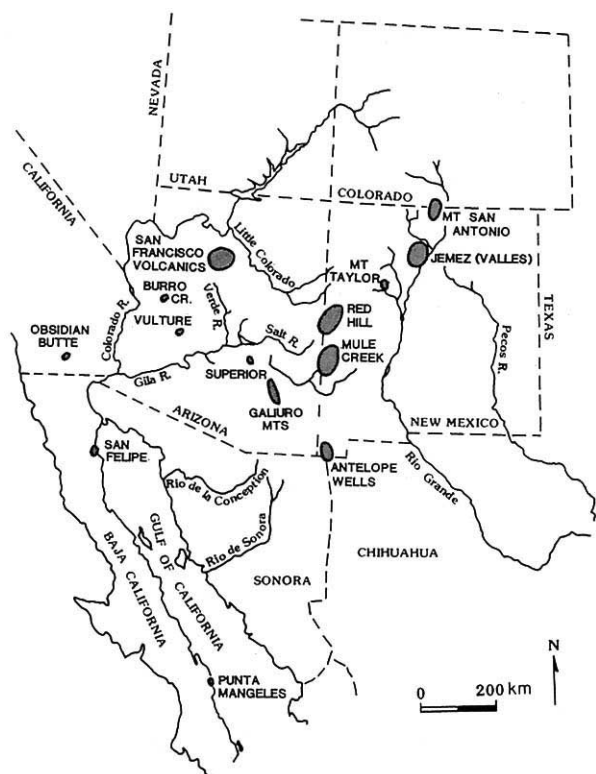


Fig. 1 Map of known obsidian sources in the southwestern United States (from Shackley⁷).

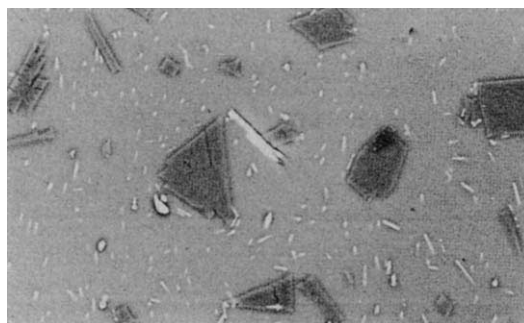


Fig. 2 BSE-image of obsidian from Government Mountain, Arizona, showing characteristic assemblage of plagioclase (dark), iron-rich biotite (light lath, centre), and iron oxides (light). Field of view, 170 µm at 20 kV.

Obsidian samples were examined by BSE from a number of sources in the southwestern United States⁷ (Fig. 1), including sources of major prehistoric importance. The initial studies reveal that obsidians typically contain large numbers of mineral inclusions. These inclusions could be imaged and identified and variations in composition, texture, and abundance, which cannot be detected optically, could be examined by BSE.

To study variability within a single source, samples were examined from Government Mountain (San Francisco Volcanic Field), which was traversed laterally and vertically. All samples of Government Mountain obsidian contain its characteristic petrographic assemblage of plagioclase, iron oxides, and iron-rich biotite (Fig. 2). Similar homogeneity was found in obsidians from other sources. Although there are variations in textures and in absolute abundances of characteristic minerals, there are not significant variations in mineral type, composition, or relative abundances for individual sources. Moreover, the obsidians were found to be mineralogically homogenous on any scale larger than a few millimetres, suggesting that it might be possible to assign sources to individual obsidian artefacts.

To study variability among different sources, obsidian sources for the San Francisco Volcanic Field of northern Arizona^{5,7} (Fig. 3) were examined. All but two were found to be petrographically unique and easily distinguished (see Table 1).

For example, Government Mountain is the only obsidian in the San Francisco Volcanic Field that contains plagioclase phenocrysts. Likewise, only the San Francisco Peaks source contains sphenes.

The two non-unique sources are Sitgreaves Peak and RS Hill. Obsidian from RS Hill, located on the north-east flank of Sitgreaves Peak, is petrographically indistinguishable from Sitgreaves Peak obsidian. Their petrographic identity and geographical proximity suggest that they, as Robinson⁵ suggested, are samples from the same magma source. On the other hand, Kendrick Peak and Slate Mountain are geographically close and indistinguishable by the trace-element data of Jack¹⁰, but can be differentiated by BSE-petrography. Government Mountain and Sitgreaves Peak are similarly indistinguishable by the trace-element data but can be differentiated using BSE. Thus, because of mineralogical complexity and homogeneity of obsidian sources, BSE petrography can not only distinguish volcanic provinces but can also identify obsidians from individual magma sources within a province, even those that cannot be differentiated by the trace-element data.

When the distinctive petrographic 'fingerprints' of the various sources were recognized, the utility of the BSE-petrographic method for provenance determination was tested. A set of obsidian flakes and points was provided by Northland Research from four loci at the 'Shelltown' Hohokam site, south-central Arizona, for source determination. Five of these artefacts had been independently assigned to known sources by the rapid-scan Rb-Sr-Zr(-Nb) XRF method of Jack and Heizer¹¹. These were examined by BSE in a blind test, randomly assigning sample numbers and using these until after source assignments were made, without knowledge of the XRF source assignments. Two of these artefacts were identified by BSE as Government Mountain obsidian; these were sourced by XRF as Government Mountain-Sitgreaves Peak, which were not resolved by XRF. The other three were identified by BSE as either Superior or Vulture sources, which were not differentiated by BSE; these were identified as Superior obsidian by XRF.

Moreover, a pattern, similar to the chronological correlation of sources discovered by Shackley¹², emerged from the BSE-results. Source assignments could be made for twelve of sixteen

Table 1 Petrographic features of the San Francisco field obsidians

Source	Feldspar cations	Common phases	Other distinctive features
Government Mountain	Ca > Na	Fe-biotite, Fe oxide	Mn-bearing pyroxene
Sitgreaves Peak	Na	Hornblende, Fe oxide	
RS Hill	Na	Hornblende, Fe oxide	
Slate Mountain	Ca > K = Na	Fe-biotite, FeTi oxide	Potassic feldspar cores
Kendrick Peak	Ca = K = Na	Fe oxide, CaFeSi-phase	
O'Leary Peak	Ca = Na = K	Fe-biotite, Fe oxide	Hyalopilitic
Robinson Crater	Na = K	Fe-biotite	Hyalopilitic
San Francisco Peak	Na	Zircon, chevkinite	Large phenocrysts

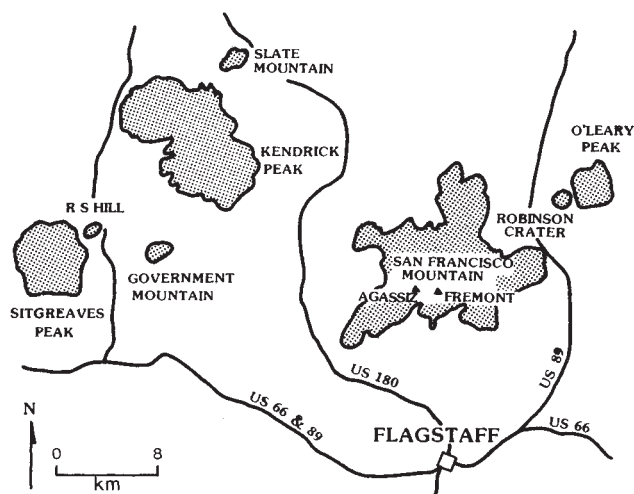


Fig. 3 Map of obsidian sources in the San Francisco Volcanic Field, Arizona (from Schreiber and Breed⁹). The location of the San Francisco Volcanics is shown in Fig. 1 (upper left).

other Shelltown specimens. Two artefacts from locus one, possibly a pre-Colonial Hohokam site, resembled source material from Red Hill, New Mexico. Artefacts from loci two and three, dated by ceramic typology as Colonial-phase Hohokam sites, matched the Government Mountain source. Artefacts from locus

four, dated by ceramic typology as a Sedentary-phase site, matched mid-Tertiary Sonoran Desert sources, mostly Vulture and Superior. Not only can individual sources be resolved but, with ceramic-chronological data, diachronic variations in procurement patterns can be detected.

Thus, BSE-petrography has the ability to uniquely characterize obsidian sources, to accurately and efficiently assign specific sources to individual obsidian artefacts, and consequently to provide valuable insights on patterns of prehistoric procurement and exchange.

We thank Steven Shackley for providing source materials for this study, Thomas Hartzel for sampling the Government Mountain source, and Northland Research for providing artefacts and ceramic-chronologic data. The work was supported by a Grant-in-Aid-of-Research from Sigma Xi to James Burton.

Received 8 September 1986; accepted 19 January 1987.

1. Cann, J. R. in *The Petrology of Archaeological Artefacts* (eds Kempe, D. R. C. & Harvey, A. P.) 227-255 (Clarendon, Oxford 1983).
2. Stevenson, D. P., Stross, F. H. & Heizer, R. F. *Archaeometry* 13, 17-25 (1971).
3. Griffin, J. B. & Gordus, A. A. *Science* 158, 382-384 (1967).
4. Robinson, B. W. & Nickel, E. H. *Am. Miner.* 64, 1322-1328 (1979).
5. Hall, M. G. & Lloyd, G. E. *Am. Miner.* 66, 362-368 (1981).
6. Krinsley, D. H., Pye, K. & Kearsley, A. T. *Geol. Mag.* 120, 109-114 (1983).
7. Shackley, M. S. paper presented at 50th Ann. Meet. Soc. Amer. Anth. (1985).
8. Robinson, H. H. *Prof. Pap. U.S. geol. Surv.* 76, (1913).
9. Schreiber, J. P. & Breed, W. J. *Plateau* 43, 115-119 (1971).
10. Jack, R. N. *Plateau* 43, 103-114 (1971).
11. Jack, R. N. & Heizer, R. F. *Contr. Univ. Calif. Arch. Res. Facility* 5, 81-95 (1968).
12. Shackley, M. S. paper presented at 51st Ann. Meet. Soc. Amer. Anth. (1986).

Liquid carbon dioxide of magmatic origin and its role in volcanic eruptions

Allan R. Chivas*, Ivan Barnes†, William C. Evans†, John E. Lupton‡ & John O. Stone§

* Research School of Earth Sciences, The Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia

† Water Resources Division, U.S. Geological Survey, Menlo Park, California 94025, USA

‡ Marine Science Institute and Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA

Natural liquid carbon dioxide is produced commercially from a 2.5-km-deep well near the 4,500-yr-old maar volcano, Mount Gambier, South Australia. The carbon dioxide has accumulated in a dome that is located on the extension of a linear chain of volcanic activity. A magmatic origin for the fluid is suggested by the geological setting, $\delta^{13}\text{C}_{\text{PDB}}$ of -4.0% for the CO_2 (where PDB represents the carbon-isotope standard), and a relatively high ^3He component of the contained helium and high $^3\text{He}/\text{C}$ ratio (6.4×10^{-10}). The $^3\text{He}/^4\text{He}$ and He/Ne ratios are 3.0 and $>1,370$ times those of air, respectively. The CO_2 , as collected at the Earth's surface at 29.5°C and 75 bar, expands more than 300-fold to form a gas at 1 atm and 22°C . We suggest that liquid CO_2 or high-density CO_2 fluid (the critical point is 31.1°C , 73.9 bar) of volcanic origin that expands explosively from shallow levels in the Earth's crust may be a major contributor to 'phreatic' volcanic eruptions and maar formation. Less violent release of magmatic CO_2 into crater lakes may cause gas bursts with equally disastrous consequences such as occurred at Lake Nyos, Cameroon, in August 1986.

The Caroline No. 1 well is situated 15 km south-east of Mt Gambier, South Australia, along the southeasterly extension of a small chain of intraplate basaltic volcanism (Fig. 1). On geomorphic grounds, the oldest volcanoes are at the north-west end of the chain. Mt Schank has an age of 4,900 yr¹ and the

Mt-Gambier composite maar^{2,3} is the youngest known volcano in Australia, with the most recent eruption only ~4,500 yr ago, according to radiocarbon dating⁴ of charcoal beneath tuff. The Caroline-well location was selected at the coincidence of a positive gravity anomaly and a domal structure indicated by a seismic reflection survey⁵. It is only speculation to suggest that these geophysical features are related to an underlying basaltic intrusion. The age of the doming is unknown. Completed in 1966 to a depth of 3.4 km, Caroline No. 1 produces (at the surface) a liquid that is ~98% CO_2 from intervals below a depth of 2.5 km where the fluid is at a temperature of 92°C and a pressure of ~240 bar. After expansion and purification, the carbon dioxide is recompressed for industrial use. Production now is 28,000 tonnes yr⁻¹. The other principal constituents of the fluid (Table 1) are methane (~1%) and nitrogen (~0.5%). Traces of ethane, helium, argon and higher hydrocarbons are present.

The CO_2 -rich fluid at Caroline No. 1 is contained in carbonate-poor sandstones interbedded with shales and siltstones of the Waarre Formation of mid-Cretaceous age⁵. Wopfner and Thornton considered⁵ that the CO_2 has a volcanic origin based on the proximity of recent volcanism and the absence of carbonate rocks both in the section containing the carbon dioxide and below. Further high-purity CO_2 (~96%) with minor methane (2.4%) and nitrogen (1.3%) was also encountered at a depth of 2.1 km in an unconformity-type trap in the now-abandoned Kalangadoo No. 1 well, 33-km north of Mt Gambier. Here, carbon dioxide occurs in fractured dolomitic sandstone of assumed Palaeozoic age just below the unconformity at the base of the Mesozoic Otway Group⁵.

The $\delta^{13}\text{C}_{\text{PDB}} (\%)$ values of -3.7 and -4.3 for the Caroline carbon dioxide are similar to those calculated for the Kilauea parental magma (-3.6 to -3.0 ; ref. 6) and slightly heavier than the mean value of about -6 that is typical of deep-seated carbon from CO_2 -bearing fluid inclusions and glasses in submarine basalts⁷⁻⁹. A range of $\delta^{13}\text{C}$ values from -9 to -2 is commonly found for deep-seated carbon^{10,11}; some south-east Australian diamonds have values as high as $+2.7$ (Craig in ref. 11; 12, 13). The $^3\text{He}/^4\text{He}$ ratio of the helium contained in the fluid from Caroline No. 1 is three times that of air and shows that even

§ Present address: Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA.

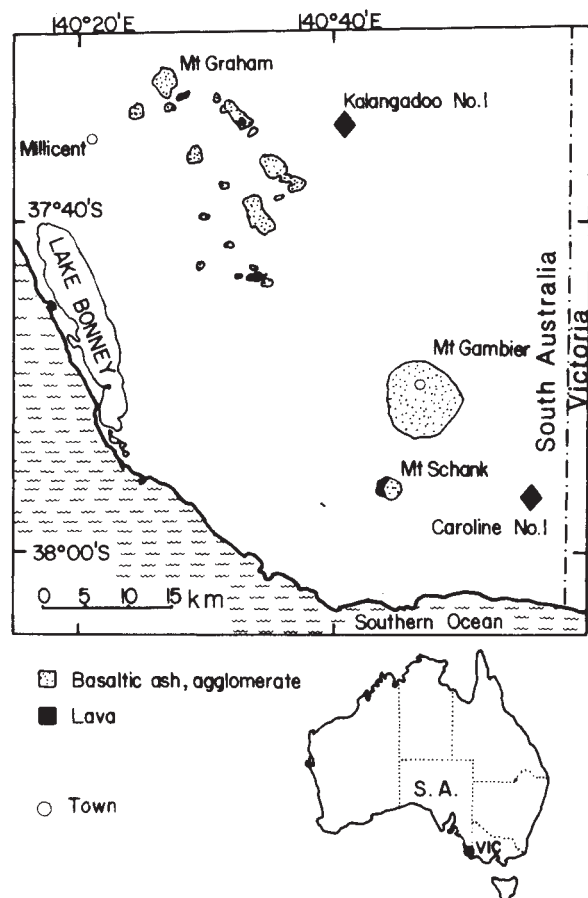


Fig. 1 Location of Caroline No. 1 and Kalangadoo No. 1 wells in relation to Quaternary volcanism, South Australia (modified from Sheard³).

though the domal structure acts as a trap for radiogenic ^4He , there is also direct contribution and collection of primordial ^3He . Both the $\delta^{13}\text{C}$ and ^3He data support a magmatic origin for the Caroline dense fluid phase. The $^3\text{He}/\text{C}$ ratio of 6.4×10^{-10} is within the range of values for fluids and tholeiitic glasses from mid-ocean-ridge basalts (MORB) of 1.4×10^{-10} – 1.8×10^{-9} (ref. 14), and higher than is typical of values for CO_2 -rich continental fluids, the majority of which are in the range 8×10^{-12} – 2×10^{-10} (see, for example, refs 15, 16, 17). The Caroline fluid has negligible ^{14}C activity, $\delta^{15}\text{N}_{\text{AN}}(\text{‰})$ of N_2 of +2.6, and $\delta^{13}\text{C}_{\text{PDB}}(\text{‰})$ of CH_4 of -40.6. The $\delta^{13}\text{C}$ value for methane is between values reported for MORB fluids (-15 to -21‰; refs 18, 19) and those typical of bacterially produced methane (-45 to -80‰, ref. 20). $\Delta^{13}\text{C}_{\text{CO}_2-\text{CH}_4}$ would indicate a temperature²¹ of 180 °C, if isotopic equilibrium were assumed. At least some of the methane, however, is likely to have an organic origin, because traces of higher hydrocarbons were present in the sampling of 22 February 1986 (0.0115% C_2H_6 , 0.0046% C_3H_8 , 0.0007% $i\text{-C}_4\text{H}_{10}$, 0.0010% $n\text{-C}_4\text{H}_{10}$, 0.0002% $i\text{-C}_5\text{H}_{12}$, 0.0002% $n\text{-C}_5\text{H}_{12}$) and approximately one barrel (0.159 m³) per year of hydrocarbon condensate is discarded from the production plant. This condensate is unusually rich in aromatic constituents, and McKirdy²² has suggested that the CO_2 has stripped small quantities of aromatic and less soluble aliphatic hydrocarbons from marginally mature, poor-quality terrestrial kerogen dispersed throughout the Waarre Sandstone. Although some of the CO_2 may also derive from organic matter, taken together, the $\delta^{13}\text{C}$ of the CO_2 , the ^3He and the $^3\text{He}/\text{C}$ ratio support a largely magmatic origin for the Caroline dense fluid phase.

Calcite ($\delta^{13}\text{C} = +6.5$) precipitated at a plumbing joint in the treatment plant is in carbon isotope equilibrium with the carbon dioxide. Baertschi²³ and Bottinga's calculations²⁴ give

Table 1 Chemical and isotope analysis of expanded fluid and precipitated CaCO_3 from Caroline No. 1 well, South Australia

		15 Nov. 1982	22 Feb. 1986
Gas analysis (vol. %)	H_2	<0.005	0.0003
	He	0.015	0.008
	O_2	<0.01	0.005
	N_2	0.74	0.39
	Ar	<0.01	0.005
	CH_4	1.57	0.90
	CO_2	97.6	98.7
	CO		<0.001
	C_2H_6	0.01	0.0115
	H_2S	<0.02	<0.001
	Total	99.94	100.02
$\delta^{15}\text{N}_{\text{AN}}(\text{‰})$ of N_2			+2.6
$\delta^{13}\text{C}_{\text{PDB}}(\text{‰})$ of CO_2		-3.7	-4.3
$\delta^{13}\text{C}_{\text{PDB}}(\text{‰})$ of CH_4			-40.6
$\delta^{18}\text{O}_{\text{SMOW}}(\text{‰})$ of CO_2		+19.1	+19.2
^{14}C activity of CO_2			<0.13 ± 0.24 c.p.m.
$\delta^{13}\text{C}_{\text{PDB}}$ of CaCO_3 precipitate			+6.5
$\delta^{18}\text{O}_{\text{SMOW}}$ of CaCO_3 precipitate			+27.9
$^3\text{He}/^4\text{He}$ (R/R_{air}) (collected 8 Aug. 1983)			3.03 ± 0.06
He/Ne (R/R_{air})			>1,370
$^3\text{He}/\text{C}$ molar ratio			6.4×10^{-10}

The CO_2 -rich samples were collected in evacuated stainless-steel cylinders (40 cm³, with 41-MPa-rated valves) directly at the wellhead at a pressure of ~7.5 MPa. The liquid CO_2 entering the evacuated vessel expanded to a dense gas (0.55 g-cm⁻³ at 22 °C) during collection. In the laboratory, the cylinder contents were slowly expanded into an evacuated glass extraction-line with attached auxiliary glass vessels having volumes of from ~6 to 17 litres. After equilibration, the expanded gases were collected in borosilicate or lead glass (for He, Ne analysis) vessels at pressures of from 1.3 to 0.4 atm. Gas analysis by gas chromatography, He abundance (for 15 Nov. 1982 only) and He isotopic composition by mass spectrometry.

AN, atmospheric nitrogen; PDB, carbon-isotope standard; SMOW, standard mean ocean water. In text, $\Delta^{13}\text{C}_{\text{A-B}} = \delta^{13}\text{C}_{\text{A}} - \delta^{13}\text{C}_{\text{B}}$, where $\delta^{13}\text{C}_{\text{PDB}}$ value for a given sample

$$= \left[\frac{(^{13}\text{C}/^{12}\text{C})_{\text{sample}} - (^{13}\text{C}/^{12}\text{C})_{\text{PDB}}}{(^{13}\text{C}/^{12}\text{C})_{\text{PDB}}} \right] 1,000$$

$\delta^{18}\text{O}_{\text{SMOW}}$ and $\delta^{15}\text{N}_{\text{AN}}$ values are based on similar expressions for $^{18}\text{O}/^{16}\text{O}$ and $^{15}\text{N}/^{14}\text{N}$ using SMOW and AN as standards, respectively.

$\Delta^{13}\text{C}_{\text{CO}_2\text{-calcite}} \sim -10$ at 25 °C; for the Caroline material we obtain $\Delta^{13}\text{C} = -10.2$. The oxygen-isotope composition of the calcite is close to that expected to precipitate from local meteoric water, from either atmospheric condensation or tap water used to hose-clean equipment.

The Caroline carbon dioxide, as collected at the surface, expands more than 300-fold to reach a pressure of 1 atm (footnote, Table 1). In volcanic areas where high-density magmatic carbon dioxide might accumulate near the Earth's surface, violent explosions could ensue if similar reservoirs of CO_2 were disrupted by some mechanism such as seismically triggered faulting or magma surges. Rapid expansion of carbon dioxide as it breached the Earth's surface need not necessarily be accompanied by igneous activity or expansion of heated groundwater as in phreatic explosions. Some outbreaks may be quite 'cold' and appear unrelated to volcanism.

As defined by Ollier², maars are landforms caused by volcanic explosion, and consist of a crater, which reaches or extends below general ground level and is considerably wider than deep, and a surrounding rim constructed of material ejected from the crater. Many maars are associated with CO_2 -rich volcanism, such as the Laacher See (carbonatite tephra) and the Eifel districts of western Germany. The Eifel is also a prolific producer of natural CO_2 gas²⁵. In the western-Victorian maar district^{2,26}, 200-km east of Mt Gambier, CO_2 -rich inclusions occur in

mantle-derived xenoliths²⁷, and carbon dioxide is reported from abandoned wells at Garvoc No. 1 and Wangoom No. 2 in that area²⁸. Curiously, the CO₂/CH₄/N₂ ratios of the gases from these wells are similar to those at Caroline No. 1 and Kalangadoo No. 1.

Numerous maars are unaccompanied by lava extrusions. Perhaps near-surface CO₂-rich explosions could explain these features. The accumulation of carbon dioxide could be either early or late in a given episode of volcanic activity. Carbon dioxide might explode in advance of the arrival of rising magma. It is equally plausible that during waning volcanism, carbon dioxide could continue to exsolve from static and solidifying magmas. Thus maar eruptions could also indicate the final stages of volcanism.

Most previous suggestions²⁹ for the mechanisms of maar formation concern interaction of magma or magmatic fluids with groundwater^{30,31} or release of CO₂ in the mantle that drives high-velocity nodule-bearing eruptions^{32,33}. Proponents advocating maar formation by magmatic gas (composition unspecified) explosions include Cloos³⁴ and Macdonald (ref. 35; pages 243, 290). Müller and Veyl describe³⁶ the formation of Nilahue maar in Chile and ascribe the regular eruptions during four days to accumulation and release of magmatic gases, thought to be mainly water vapour. Our results suggest that large volumes of high-density magmatic carbon dioxide may accumulate at relatively shallow depth in regions of basaltic volcanism. Gas release could be massive, gradual or episodic, so that single-event, protracted-multiple or rejuvenated maars may be formed as have been described². Of particular significance is the association of CO₂-rich discharges both before and after explosive formation of the Ukinrek maars in Alaska³⁷. More gradual accumulation of volcanic CO₂ in maar lakes may have been ultimately responsible for the lethal gas bursts from lakes Monoun³⁸ (15 August 1984) and Nyos³⁹⁻⁴² (21 August 1986) in Cameroon.

We thank T. L. Dowling, E. Laing, C. J. Radnell and C. M. Vlastuin for their assistance in the laboratory, and P. De Deckker for fieldwork. We thank H. Craig and R. Poreda for access to the isotope laboratory, Scripps Institution of Oceanography. We thank the staff of Liquid Air Australia, Alliance Oil Development Australia, and SANTOS for assistance and access to the Caroline No. 1 plant and J. Cowley, J. Stringer and M. Stevenson for preparation of the manuscript. Constructive reviews were provided by D. K. Bailey and R. A. Exley.

Received 22 October; accepted 31 December 1986.

- Smith, B. W. & Prescott, J. *Aust. J. Earth Sci.* (in the press).
- Ollier, C. D. *Bull. volcan.* **31**, 45-73 (1967).
- Sheard, M. J. *Trans. R. Soc. S. Aust.* **102**, 125-139 (1978).
- Blackburn, G., Allison, G. B. & Leaney, F. W. J. *Trans. R. Soc. S. Aust.* **106**, 163-167 (1982).
- Wopfner, H. & Thornton, R. C. N. in *The Otway Basin of Southeastern Australia* (eds Wopfner, H. & Douglas, J. G.) 377-384 (Spec. Bull. geol. surveys of South Australia and Victoria, 1971).
- Gerlach, T. M. & Thomas, D. M. *Nature* **319**, 480-483 (1986).
- Pineau, F., Javoy, M. & Bottinga, Y. *Earth planet. Sci. Lett.* **29**, 413-421 (1976).
- Moore, J. G., Batchelder, J. N. & Cunningham, C. G. *J. Volcan. geotherm. Res.* **2**, 309-327 (1977).
- Exley, R. A., Mathey, D. P., Clague, D. A. & Pillinger, C. T. *Earth planet. Sci. Lett.* **78**, 189-199 (1986).
- Craig, H. *Geochim. cosmochim. Acta* **3**, 53-92 (1953).
- Wickman, F. E. *Geochim. cosmochim. Acta* **9**, 136-153 (1956).
- Sobolev, N. V., Yefimova, E. S., Lavrent'yev, Y. G. & Sobolev, V. S. *Dokl. (Proc.) Acad. Sci. U.S.S.R.* **274**, 148-153 (1985) (translated from *Dokl. Akad. Nauk SSSR*, 1984).
- Galimov, E. M. *Geochem. Int.* **22**(1), 118-142 (1985).
- Des Marais, D. J. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 602-611 (Geophys. Monogr. 32, Am. Geophys. Union, Washington, 1985).
- Kamenskiy, I. L., Lobkov, V. A., Prasolov, E. M., Beskrovnyy, N. S., Kudryavtseva, E. I., Anufriyev, G. S. & Pavlov, V. P. *Geochem. Int.* **13**(3), 35-48 (1976).
- Sano, Y. & Wakita, H. *J. geophys. Res.* **90**, 8729-8741 (1985).
- Torgersen, T., Lupton, J. E., Sheppard, D. S. & Giggenbach, W. F. *J. volcan. geotherm. Res.* **12**, 283-298 (1982).
- Welhan, J. A. & Craig, H. in *Hydrothermal Processes at Seafloor Spreading Centers* (eds Rona, P. A. et al.) 391-409 (Plenum, New York, 1983).
- Kim, K.-R., Welhan, J. A. & Craig, H. *Trans. Am. geophys. Un. (Eos)* **65**, 973 (1984).
- Deines, P. in *Handbook of Environmental Isotope Geochemistry I. The Terrestrial Environment* (eds Fritz, P. & Fontes, J. C.) 329-406 (Elsevier, Amsterdam, 1980).
- Bottinga, Y. *Geochim. cosmochim. Acta* **33**, 49-64 (1969).
- McKirdy, D. *Rep. Aust. miner. Dev. Lab. (AMDEL) F 6433/86* (AMDEL, Adelaide, 1986).

- Baertschi, P. *Miner. petrogr. Mitt.* **37**, 73-152 (1957).
- Bottinga, Y. *J. phys. Chem.* **72**, 800-808 (1968).
- Puchelt, H. in *Plateau Uplift. The Rhenish Shield—A Case History* (eds Fuchs, K., von Gehlen, K., Mälzer, H., Murawski, H. & Semmel, A.) 152 (Springer, Berlin, 1983).
- Ollier, C. D. & Joyce, E. B. *Proc. R. Soc. Vict.* **77**, 357-376 (1964).
- Andersen, T., O'Reilly, S. Y. & Griffin, W. L. *Contr. Miner. Pet.* **88**, 72-85 (1984).
- Wopfner, H., Kenley, P. R. & Thornton, R. C. N. in *The Otway Basin of Southeastern Australia* (eds Wopfner, H. & Douglas, J. G.) 385-435 (Spec. Bull. geol. surveys of South Australia and Victoria, 1971).
- Ollier, C. D. in *Physical Volcanology* (eds Civetta, L., Gasparini, P., Luongo, G. & Rapolla, A.) 289-310 (Elsevier, Amsterdam, 1974).
- Lorenz, V. in *Kimberlites. I: Kimberlites and Related Rocks* (ed. Kornprobst, J.) 299-307 (Elsevier, Amsterdam, 1984).
- Lorenz, V. *Trans. geol. Soc. S. Afr.* **88**, 459-470 (1985).
- Bailey, D. K. *Phil. Trans. R. Soc. A* **297**, 309-322 (1980).
- Bailey, D. K. *Trans. geol. Soc. S. Afr.* **88**, 449-457 (1985).
- Cloos, H. *Geol. Rdsch.* **32**, 703-800 (1941).
- Macdonald, G. A. *Volcanoes* (Prentice-Hall, New Jersey, 1972).
- Müller, G. & Veyl, G. *20th Int. Geol. Congr. (Mexico)* (ed. García, A.) **1**(2), 375-396 (Editorial Stylo, Mexico City, 1957).
- Barnes, I. & McCoy, G. A. *Geology* **7**, 434-435 (1979).
- Sigurdsson, H., Devine, J. D., Tchoua, F. M., Presser, T. S., Pringle, M. K. W. & Evans, W. C. *J. Volcan. geotherm. Res.* **31**, 1-16 (1987).
- Katzoff, J. A. *Trans. Am. geophys. Un. (Eos)* **67**, 689 (1986).
- Scientific Event Alert Network (SEAN) Bull. **11**, 2-5 (Smithsonian Institution, Washington, 1986).
- Freeth, S. J. & Kay, R. L. F. *Nature* **325**, 104-105 (1987).
- Tuttle, M. L. et al. *U.S. Geol. Survey Open-File Report* 87-97 (1987).

A spreading episode at the southern end of the San Andreas fault system

Paul G. Silver & J. N. Valette-Silver

Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington DC 20015, USA

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, USA

The Cerro Prieto region of Baja California, Mexico is one of the few locations in the world where on-land spreading is indicated¹, and represents a transition between the East Pacific Rise and the strike-slip San Andreas fault system (Fig. 1). It is characterized by high heat flow and volcanic activity, and contains one of the world's most closely monitored geothermal fields, whose heat source is thought to be a magmatic body emplaced at 5-6 km depth². Cerro Prieto is thus an excellent area in which to examine the spatial and temporal properties of the spreading process in detail. Here we present four lines of evidence that an on-land spreading episode occurred at Cerro Prieto between 1979 and 1981: the occurrence of the Imperial Valley earthquake of 15 October 1979 and the Victoria earthquake of 9 June 1980, both of which produced slip near Cerro Prieto; extension in the expected NW-SE spreading direction, both from direct measurements of areal strain and from the combined slip on the two faults; subsidence of up to 25 cm near the Cerro Prieto geothermal field, from levelling and gravity data; and an abrupt rise in temperature and HCO₃⁻ concentration in the Cerro Prieto geothermal reservoir, both of which indicate the presence of magma at depth³.

Earthquakes. Directly north and south of Cerro Prieto (Fig. 1) are the strike-slip Imperial and Cerro Prieto faults. In an eight-month period, both faults failed, producing the Imperial Valley earthquake of 15 October 1979 ($M_s = 6.9$) and the Victoria earthquake of 9 June 1980 ($M_s = 6.4$). Most of the research on the Imperial Valley earthquake has focused on the rupture north of the California-Mexico border, where all the surface slip and most of the aftershocks were found. However, recent studies^{4,5} indicate significant rupture to the south-east as well, towards Cerro Prieto. There was also a cluster of aftershocks located at the southern end of the Imperial fault (Fig. 1), indicating that either co-seismic or post-seismic slip took place there. The Victoria event appeared to rupture as far north as Cerro Prieto; several normal-faulting aftershocks were located in the proposed spreading area⁴ (Fig. 1).

Extension. Simply by the geometry of the Cerro Prieto region, the ~0.5 m of the earthquake-related right-lateral slip on both

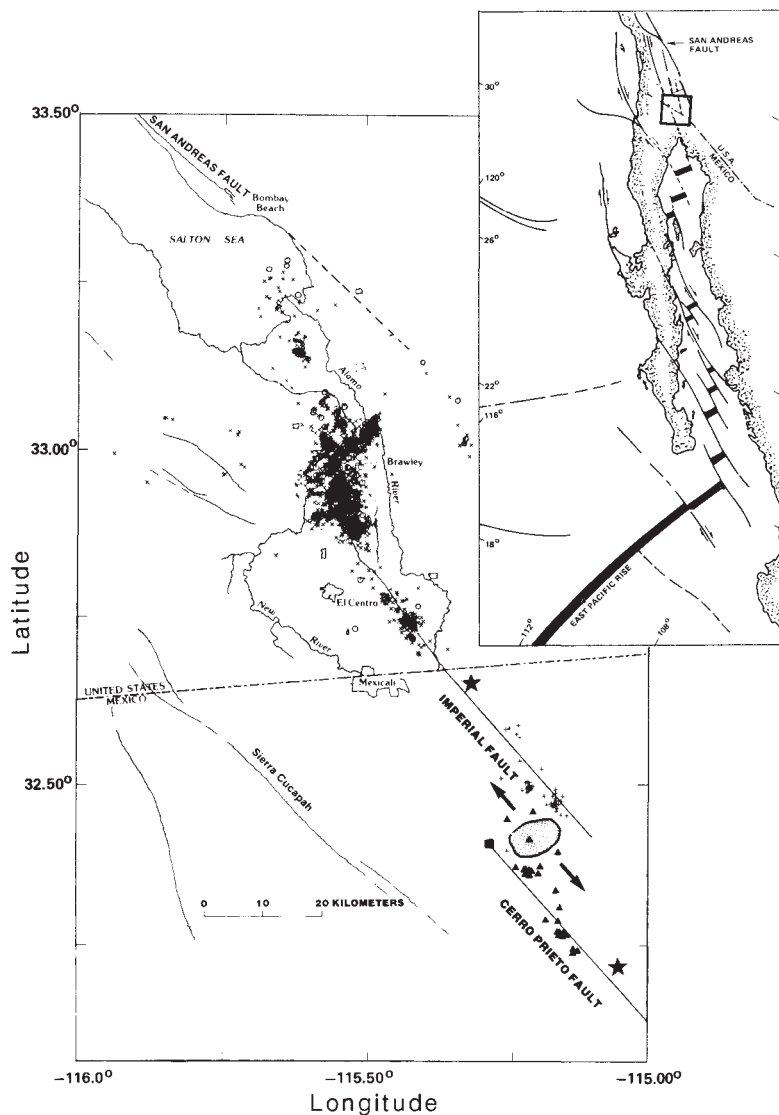


Fig. 1 Inset: tectonic setting of Cerro Prieto study area (box). This region is a location of on-land spreading and represents the transition between the San Andreas fault system to the north and East Pacific Rise to the south. The Rise system, represented by the heavy bars, intersects the continent at the top of the Gulf of California. Enlargement: Cerro Prieto study area is between the Cerro Prieto and Imperial faults; the grey region denotes the Cerro Prieto geothermal field, with the arrows giving the expected direction of spreading. The star on the Imperial fault marks the epicentral location of the 15 October, 1979 Imperial Valley earthquake ($M_S = 6.9$) with the aftershocks given by crosses, circles and xs. The star on the Cerro Prieto fault marks epicentral location of 9 June, 1980 Victoria earthquake ($M_S = 6.4$) with the aftershocks given by triangles. Square denotes the Cerro Prieto Volcano. Figure adapted from Silver and Masuda⁴.

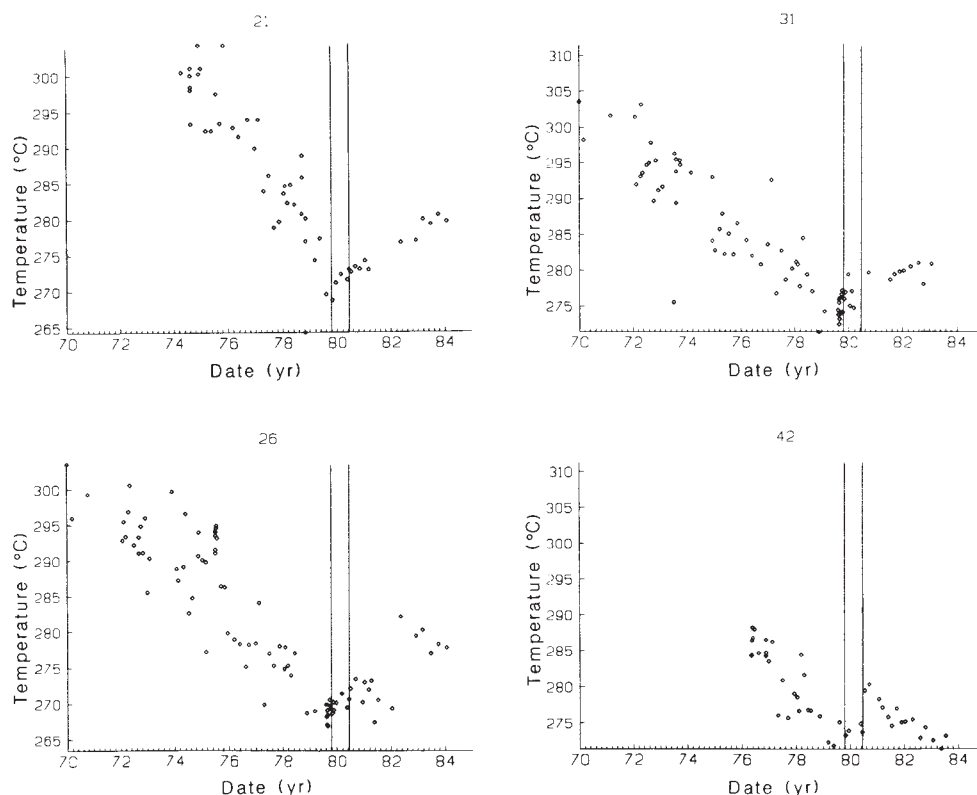
the Imperial and Cerro Prieto faults⁴ suggests that about 0.5 m of extension occurred in this region. In addition, Lisowski and Prescott⁶ performed a geodetic survey of horizontal deformation in the Cerro Prieto area using a regional trilateration network. Restricting the stations to those contained in the proposed spreading region, and for the time period covering both the Imperial Valley and Victoria earthquakes, the resulting surface strain tensor yields overall dilatation with a principal extensional strain of $2-3 \times 10^{-6}$ trending $N43^\circ W$ (W. H. Prescott, personal communication), which is the expected spreading direction. Simple modelling of a 0.5 m extensional event at a depth of 5 km produces a strain of comparable magnitude.

Subsidence. Information on subsidence comes from levelling data^{7,8} and repeated gravity measurements⁹. The levelling data indicate up to 0.25 m of subsidence in the vicinity of the geothermal field¹⁰. The gravity measurements show significant positive gravity changes for the period of time 1978–1983 with most of this increase, 80 μgal , occurring during the interval 1980–1981 which includes the Victoria event. If we attribute the gravity signal to an elevation change, this corresponds to about 0.25 m of subsidence, consistent with the levelling data.

Magma. For Cerro Prieto, the direct observation of magma associated with a rifting episode (such as in Iceland¹¹ and the Afar, Africa¹²) would most certainly be obscured by the almost 5 km of sediments overlying the basement. However, magma involvement may still be detected through fluctuations in temperature and chemistry of a geothermal reservoir. To evaluate this possibility, we have examined the data collected over the

last fifteen years to monitor the Cerro Prieto reservoir (C.F.E. A. H. Truesdell and C. Janick, personal communication). Data from 16 wells with measurements before 1977 were chosen, allowing for the establishment of a stable baseline of chemical behaviour. For each well, the reservoir temperature T was calculated using a particularly reliable geothermometer based on the concentration of Ca, Na and K in water¹³. Figure 2 shows plots of T versus time for some of the wells, along with the origin times of the Imperial Valley and Victoria events. Before the earthquakes, the temperature generally decreases with time, presumably due to production. At about the time of the Imperial Valley event, however, there is an interruption of this trend and a rise in temperature. We can quantify this change by estimating the slope $r = dT/dt$ of a regression line, for the interval before 1979.5, r_1 , and comparing it to the period during the proposed anomaly, 1979.5–1981.0, r_2 . For some wells, the first year of production was not used, because there are often large temperature excursions before equilibrium is established; some extreme values of temperature, considered as unreliable, were excluded as well. The estimated slopes r_1 and r_2 along with 1σ uncertainties are shown in Fig. 3. The most striking feature of the plot is that in all cases, r_2 is more positive than r_1 . The average values for all wells are $\bar{r}_1 = -1.70 \pm 0.13^\circ \text{C yr}^{-1}$ and $\bar{r}_2 = +2.58 \pm 0.26^\circ \text{C yr}^{-1}$ which are significantly different at the 99% confidence level. The thermal anomaly, defined as $r_2 - r_1$ has good geographical coherence, producing a NW–SE trending linear feature (Fig. 4a). This pattern is similar to that found for the measured isotherms of the field at 1,000 m depth, close to

Fig. 2 Temperature in °C using a Na-K-Ca geothermometer for wells 21, 26, 31 and 42. The horizontal axis shows time in years. The two vertical lines give the origin times of the Imperial Valley earthquake of 15 October 1979 and the Victoria earthquake of 9 June 1980. Location of wells is shown in Fig. 4.



the production depth of most of the wells¹⁴ (Fig. 4b), indicating that the anomaly is strongest in the hottest part of the reservoir.

Another important indicator of magmatic influence on reservoir chemistry is the concentration of CO₂. In the case of the Krafla geothermal field in Iceland, the interaction of magma with the reservoir produced, within only a few weeks, a significant increase in the CO₂ concentration^{3,15} that lasted more than two years (H. Kristmannsdottir, personal communication). For Cerro Prieto, we examined the ion HCO₃⁻, which reflects the dissolution of CO₂ in geothermal waters, and found that most wells indeed showed a dramatic increase of a factor of two to three in concentration between 1979.5 and 1981.0. A regression analysis of these data yields $\bar{r}_1 = +3.8 \pm 1.3$ p.p.m. yr⁻¹ and $\bar{r}_2 = +39.4 \pm 4.1$ p.p.m. yr⁻¹.

The data presented above suggest that the occurrence of two moderate-sized earthquakes, crustal deformation in the region between these events, and a thermal/chemical anomaly in the Cerro Prieto geothermal reservoir, represent different manifestations of a tectonic event, most simply interpreted as a spreading episode.

One could perhaps object to the above interpretation by arguing that the close spatial and temporal proximity of the two earthquakes is merely coincidental and that the other indicators are the result of geothermal production rather than regional tectonics. In the case of the horizontal deformation however, Lisowski and Prescott⁶ concluded that the observed strain for 1979–1981 was dominated by the occurrence of the two earthquakes and apparently masked any production related changes. For the levelling and gravity data, it might be possible to produce subsidence of the magnitude discussed here, for example, by the response to withdrawal of fluid. However, Grannel *et al.*⁹ have noted that referenced to Sierra Cucapah, west of the Cerro Prieto area (Fig. 1), there was a significant gravity decrease of 20 μ gal over the Cerro Prieto volcano, corresponding to at least 5 cm of uplift, during the same time period that maximum subsidence near the geothermal field was observed. It would be difficult to attribute both the uplift and subsidence to local reservoir conditions and we conclude that there must be a significant tectonic component in the data.

Although a rise in temperature and in HCO₃⁻ concentration between 1979.5 and 1981.0 appears to be well substantiated by the data, there exists the possibility of production-induced changes. To explain the geographical coherence of the observed anomaly, the entire reservoir must be affected, rather than the local conditions around individual wells. However, the two procedures which might have produced a reservoir-wide effect, doubling production in 1979 and performing a reinjection test between 1979 and 1980, are expected to produce a temperature decrease^{16,17} and thus cannot easily explain the observed rise. Furthermore, it appears difficult to produce such a dramatic rise in HCO₃⁻ concentration by these procedures. Although it would be impossible to rule out categorically manmade causes of the anomaly, we conclude that a tectonic origin is more likely.

The timing and magnitude of the thermal anomaly can be plausibly explained by the intrusion of magma ($T > 1,000$ °C) at the level of the basement (~ 5 km). Flow of water along

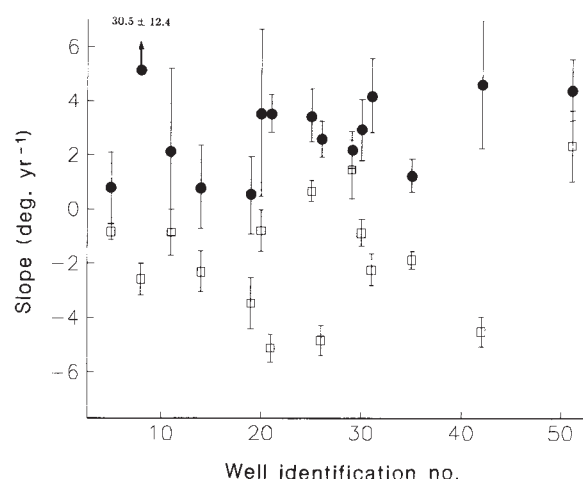


Fig. 3 Estimated temperature changes in °C yr⁻¹ for the periods 1979.5–1981.0 (solid circles) and before 1979.5 (open squares). Error bars are 1 σ uncertainties.

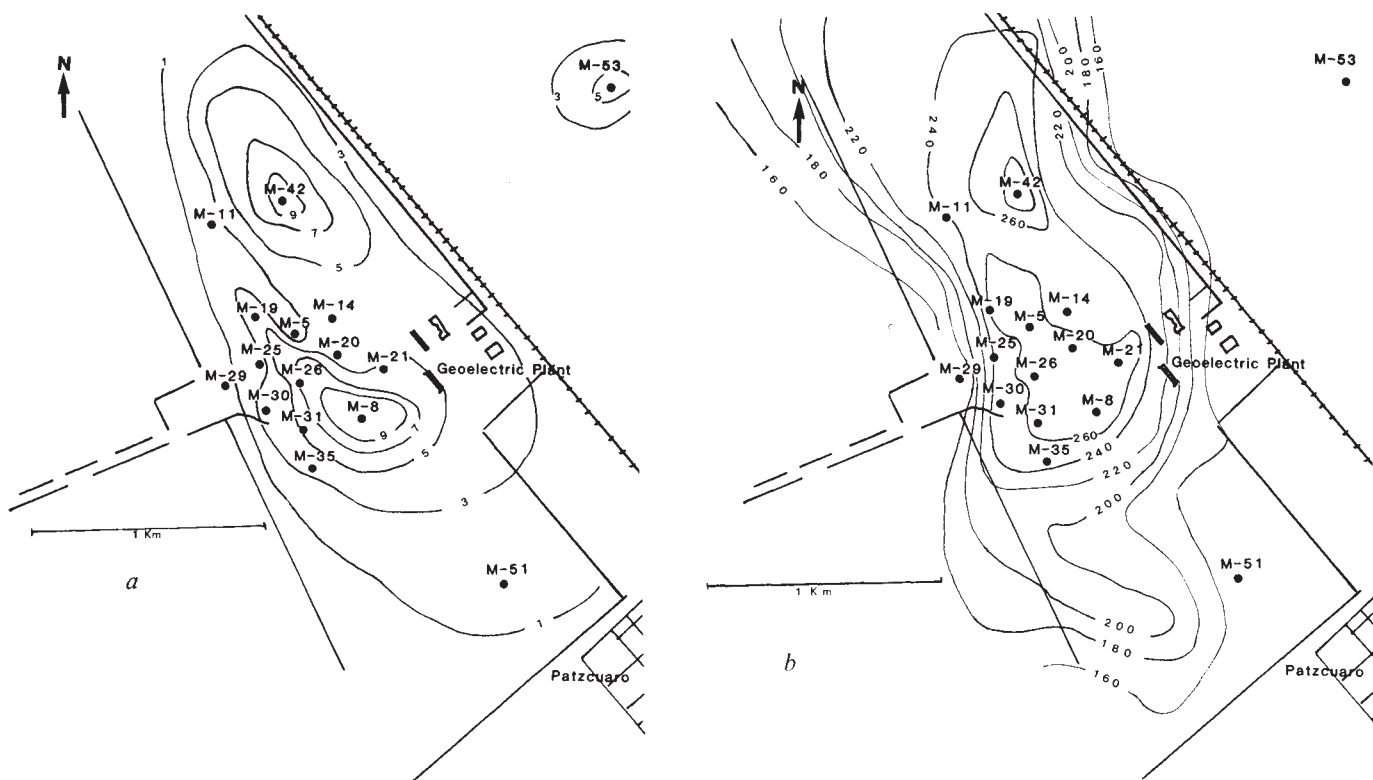


Fig. 4 *a*, Contour plot of the thermal anomaly measured by $\delta T = T_2 - T_1$ in $^{\circ}\text{C}$ (see text) for the wells studied. The numbers preceded by an M are the well identification numbers. *b*, Isotherms at 1,000 m depth in $^{\circ}\text{C}$. Note close resemblance of the two contour plots.

fractures in geothermal systems is known to occur at velocities of a few km per day¹⁸. This is sufficiently fast to associate the observed temperature rise with the vertical ascent of hot water from a basement-level magmatic event at about the time of the Imperial Valley earthquake. The Cerro Prieto field has several large vertical faults and is at least partially fracture-controlled¹⁹, if not fracture-dominated.

The volume of hot water required to produce the observed thermal anomaly of $4.3^{\circ}\text{C yr}^{-1}$ is estimated to be $200 \text{ m}^3 \text{ h}^{-1}$, which represents only $\sim 5\%$ of the production flow rate²⁰ (this assumes the affected reservoir volume is 2 km by 1 km (Fig. 4) by 0.2 km in thickness¹⁴, the porosity is 10%, the reservoir and ascending water temperatures are 300°C and 400°C respectively, and ignores the slow, conductive heating of the sedimentary matrix). This would be expected to produce only a subtle rise in reservoir level (and well-head pressure), which is indeed revealed by careful analysis of the data.

One of the intriguing consequences of a spreading-episode interpretation is the presence of a causal link between the two earthquakes. A plausible mechanism for producing such a link is strain migration, where the strain field produced by one earthquake spreads to an adjacent fault, making slip more probable on the second fault. One can calculate an apparent strain migration velocity v_a by taking the ratio of the distance to the time-separation of the two events; for the events under study, v_a is about 100 m per day. This number seems reasonable from a rheological point of view. For a simple model consisting of an elastic layer over a viscoelastic halfspace, the velocity v_s at which strain migration would be expected to occur, can be formed from the ratio of the elastic layer thickness (lithosphere) L and the relaxation time τ , (J . Rundle, personal communication), where $\tau = 2\nu/\mu$, μ is the shear modulus, and ν is the asthenosphere viscosity. Taking plausible values of $\mu = 7 \times 10^{10} \text{ Pa}$, $\nu = 10^{18} \text{ Pa s}$, and $L = 40 \text{ km}$, we obtain $v_s \sim 100 \text{ m per day}$ suggesting that the apparent strain migration may, in fact, be real. The Cerro Prieto and Imperial faults appear to have been coupled in the past. In 1915, there was a moderate-

sized earthquake on the Imperial fault followed by one on the Cerro Prieto fault six months later²¹. This second event appeared to be accompanied by a geyser-like eruption near the location of the present geothermal field.

We conclude that the data set presented above represents various manifestations of a spreading episode. We infer that the beginning of spreading approximately coincided with the occurrence of the Imperial Valley earthquake, producing an episode of extension and subsidence in the Cerro Prieto area, and the movement of magma at depth whose thermal/chemical signature quickly made its way to the surface by the upward movement of heated water through fractures. The tectonic activity appears to have subsided after the occurrence of the Victoria Earthquake. There appears to be a causal relationship between the two earthquakes and the thermal/chemical anomaly in the reservoir. This could possibly be exploited in the search for earthquake precursors. Such a relationship need not be confined to spreading environments; subduction-related volcanism often produces geothermal reservoirs in close proximity to seismogenic trench environments. The recent Michoacan, Mexico earthquake of 1985, for example, appears to have had a geothermal precursor²².

The proposed tectonic episode may not have been restricted to the region discussed in this report. Of particular interest is the adjacent area comprising the Imperial fault, the southern end of the San Andreas fault, and the intervening Brawley-Salton Sea area which is morphologically similar to the region studied here (Fig. 1). There appears to have been a strain event along the southern end of the San Andreas fault between 1978 and 1979²³ as well as a large number of aftershocks in the Brawley area as a result of the Imperial Valley earthquake. These phenomena may represent a similar and related kind of event.

We thank the Comision Federal de Electricidad, Coordinadora Ejecutiva de Cerro Prieto for their help in gathering the data, in particular Ing. A. Manóñ, R. Marquez, F. Bermejo, B. Terrazas as well as A. Truesdell and C. Janick from the US Geological Survey. We also thank W. Elders, P. Kenyon, J. Rundle, S. Greer, and M. Sorrey for useful discussions.

Received 12 September 1986; accepted 11 February 1987.

- Elders, W. A., Rex, R. W., Meidav, T., Robinson, P. T. & Biehler, S. *Science* **178**, 15-24 (1972).
- Elders, W. A., Bird, D. K., Williams, A. E. & Schiffman, P. *Geothermics* **13**, 27-47 (1984).
- Amorsson, S. *J. geophys. Res.* **91**, 261-12,268 (1986).
- Silver, P. G. & Masuda, T. *J. geophys. Res.* **90**, 7639-7651 (1985).
- Anderson, J. G. & Silver, P. G. *Geophys. Res. Lett.* **12**, 349-352 (1985).
- Lisowski, M. & Prescott, W. H. in *Proc. Fourth Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, Guadalajara, Jalisco (eds Mañón A. & Lippmann, M.) (Coordinadora Ejecutiva de Cerro Prieto, Mexicali, B.C. Mexico, 1982).
- Peña de la, A. L. in *Proc. Third Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, San Francisco, California (eds Mañón, A. & Lippmann, M.) (Lawrence Berkeley Laboratory, Berkeley, California, USA, 1981).
- Massey, B. L. in *Proc. Third Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, San Francisco, California (eds Mañón, A. & Lippmann, M.) 306-307 (Lawrence Berkeley Laboratory, Berkeley, California, USA, 1981).
- Grannell, R. B., Zhou, H. & Wyman, R. M. *Geotherm. Resources Council Transactions* **8**, 203-206 (1984).
- Grannell, R. B., Wyman, R. M., Leggewie, R. M. & Randall, G. A. in *Proc. Fourth Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, Guadalajara, Jalisco 61 (eds Mañón, A. & Lippmann, M.) (Coordinadora Ejecutiva de Cerro Prieto, Mexicali, B.C. Mexico, 1982).
- Björnsson, A., Johnsen, G., Sigurdsson, S., Thornbergsson, G. & Tryggvason, E. *J. geophys. Res.* **84**, 3029-3038 (1979).
- Ruegg, J. C., Lepine, J. C., Tarantola, A. & Kasser, M. *Geophys. Res. Lett.* **6**, 817-820 (1979).
- Fournier, R. O. & Truesdell, A. H. *Geochim. cosmochim. Acta* **37**, 1255-1275 (1973).
- Peña de la, A. L., Puente, C. I. & Diaz, C. E. in *Proc. Second Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, Mexicali, BC (eds Mañón, A. & Lippmann, M.) 29-56 (Coordinadora Ejecutiva de Cerro Prieto, Mexicali, B.C. Mexico, 1979).
- Armansson, H., Gislason, G. & Hauksson, T. *Geochim. cosmochim. Acta* **46**, 167-177 (1982).
- Truesdell, A. H., D'Amore, F. & Nieva, D. *Geothermal Resources Council Trans.* **8**, 223-229 (1984).
- Grant, M. A., Truesdell, A. H. & Mañón, A. M. *Geothermics* **13**, nos 1-2, 117-140 (1984).
- Donaldson, I. G. A. *Rev. Earth planet. Sci.* **10**, 377-395 (1982).
- Hallman, S. E., Lippmann, M. J., Zelwer, R. & Howard, J. H. *Bull. Am. Assoc. Petrol. Geol.* **68**, 18-30 (1984).
- Goyal, K. P., Miller, C. W., Lippmann, M. J. & Vander Haar, S. P. in *Proc. Third Symp. on the Cerro Prieto Geothermal Field, B.C. Mexico*, San Francisco, California (eds Mañón, A. & Lippmann, M.) 496-500 (Lawrence Berkeley Laboratory, Berkeley, California, USA, 1981).
- Bull. seism. Soc. Am.* **6**, 181-184 (1916).
- Santoyo, E., Verma, S. P., Nieva, D. & Portugal-Marín, E. *Bull. Volcanol.* (in the press).
- Savage, J. C., Prescott, W. H., Lisowski, M. & King, N. E. *J. geophys. Res.* **86**, 6991-7001 (1981).

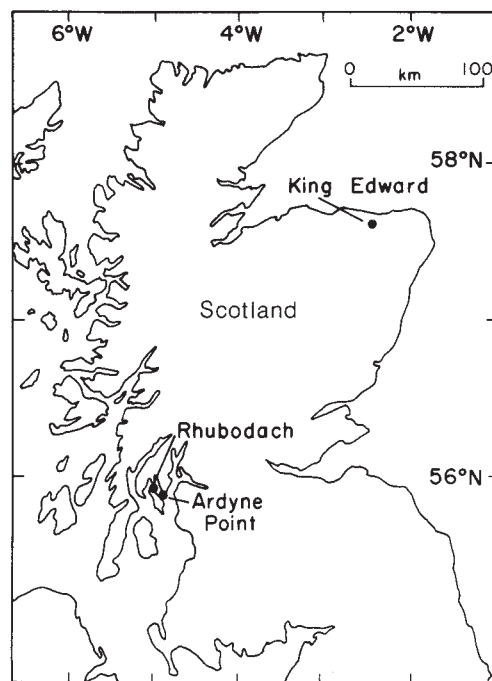


Fig. 1 Map showing the locations mentioned in the text.

Racemization-derived late Devensian temperature reduction in Scotland

G. H. Miller*, A. J. T. Jull†, T. Linick‡, D. Sutherland‡, H. P. Sejrup§, J. K. Brigham||, D. Q. Bowen¶ & J. Mangerud§

* INSTAAR and Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309, USA

† TAMS Radiocarbon Facility, Department of Physics, University of Arizona, Tucson, Arizona 85721, USA

‡ Department of Geography, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

§ Geological Institute, Division B, University of Bergen, 5000 Bergen, Norway

|| Department of Geology, University of Alberta, Edmonton, Alberta, Canada T6G 2E3

¶ Department of Geography, Royal Holloway and Bedford New College, Egham Hill, Egham, Surrey TW20 0EX, UK

We report here that shells of the marine bivalve mollusc *Arctica islandica*, collected near King Edward, Scotland, yield isoleucine epimerization D/L ratios only slightly higher than ratios in late-glacial specimens of the same taxon collected from nearby sites. We have determined the ages of two of the King Edward shells by tandem-accelerator mass spectrometry (TAMS) radiocarbon dating to be >40 kyr. From the difference in D/L ratios, we calculate the average temperature depression during the last glaciation to have been $\geq 9^\circ\text{C}$ below Holocene temperatures.

The extent of amino acid racemization (or epimerization) in proteinaceous residues has become a standard Quaternary dating tool applied primarily to molluscan and foraminiferal fossils. The primary rate-controlling parameters are temperature and, to a lesser extent, taxonomy. Recent summaries include applications to high-sea-level events along north-west Europe, the UK, the Mediterranean basin, and the USA¹⁻⁴. The alterna-

tive application of the method, palaeothermometry, has received less attention, even though such calculations can be made with considerably greater precision than absolute dates⁵. Here we use the difference in amino acid D/L ratios in dated shells to provide minimum estimates for the temperature depression that occurred over Scotland during the middle and late Devensian.

At King Edward in north-east Scotland (Fig. 1) Jamieson⁶ described an apparently *in situ* marine shell bed overlain by a till (which was shelly in its lower portion) and stratified sands and gravels. The stratigraphy and interpretation was confirmed by Horne⁷. Jamieson's section is no longer exposed, but some 200 m to the north-east a till, similar to that described by Jamieson and containing abundant shell fragments, occurs by the King Edward Burn. Samples of the marine pelecypod *Arctica islandica* collected from the till were used for the amino acid and radiocarbon analyses. We consider the shells to be derived from the nearby shell bed, thought to be the product of a marine invasion of the area before the last glacial advance in this part of north-east Scotland.

Protein involved in the secretion of molluscan shells is inter-layered within the carbonate matrix and thereby protected from biological attack after the death of the organism. The degradation of indigenous proteinaceous residues proceeds slowly over geological time at a rate controlled by ambient temperature. We monitor degradation by measuring the extent of isoleucine epimerization. In a modern shell, essentially all of the isoleucine is in the L-configuration. With increasing age the proportion of the D-configuration increases until an equilibrium mixture of D- and L-amino acids is attained (~ 2 Myr at 10°C); a summary of the controls on isoleucine epimerization is given in ref. 1.

D/L ratios in samples from independently dated horizons can be used to calculate the difference in the average temperature between dated levels. Unlike floral and faunal assemblages that provide instantaneous proxy data on palaeoclimates, amino acid palaeothermometry provides an integrated record of thermal conditions since the mollusc died.

The extent of isoleucine epimerization was measured in *Arctica islandica* from the King Edward locality at three laboratories (Table 1). Although Sejrup⁸ has found significant intra-shell variation in amino acid content and D/L ratio in *Arctica* from Norway, our measurements were made exclusively on the dense middle layer of the shell, and the similarity in ratios both within

Table 1 D/L ratios in *Arctica islandica* from localities in Scotland

Locality	Laboratory identification*		D/L(aIle/Ile)
King Edward	AAL-4361	A	0.078
		B	0.079
King Edward	BAL-327	A	0.080
		B	0.080
		C	0.058
King Edward	AAL-4434	A	0.095
		B	0.079
		C	0.073
King Edward	ABER-379	A	0.082
		Mean:	0.078 ± 0.010
Ardyne	AAL-2826	A	0.047
Elf Pit		B	0.053
Ardyne	AAL-2827	A	0.053
Shell Pit		B	0.061
Rhubodach	BAL-539	A	0.062
		B	0.056
		Mean:	0.055 ± 0.006

* AAL, University of Colorado, USA; BAL, University of Bergen, Norway; ABER, University of Wales, Aberystwyth, UK.

and between laboratories argues that this variability is insignificant in our samples. The mean D/L ratio based on nine separate preparations is 0.078 ± 0.010 . This ratio is substantially below that found in last-interglacial *Arctica* from north-west Germany and Denmark (0.17 ± 0.02 , $n = 42$; ref. 1), but is only slightly higher than the D/L ratios in dated late-glacial interstadial *Arctica* from sites in the Firth of Clyde, western Scotland, that are radiocarbon dated between 11 and 11.5 kyr BP. Four individual *Arctica* analysed from these sites and two additional individuals from the Rhubodach locality 9 km to the west gave a mean D/L ratio of 0.055 ± 0.06 (Table 1). The small difference in D/L ratio between the late-glacial and King Edward shells implies either that they are of similar age or that the intervening period was much colder than the period since 11 kyr BP, thereby decreasing the epimerization rate in the older shells. To test these two options, we used the TAMS radiocarbon facility at the University of Arizona to date divided samples of two shells from King Edward that we had used for the amino acid study.

Two shells, AAL-4361 and AAL-4434, were selected for accelerator dating. One quarter of each fragment was dissolved in purified 2 M HCl before submission. Carbon dioxide was produced from ~30 mg of each shell by hydrolysis in 85% H_3PO_4 . The CO_2 was reduced to CO over hot Zn at 400 °C, and further to graphite over Fe at 750 °C in the presence of hydrogen (ref. 9). The graphite powder produced was packed into a target holder and mounted in the accelerator ion source. Carbon-14 was measured both in samples and in two standards (NBS oxalic acid I and II); complete details of the accelerator ^{14}C measurement procedures are given by Linick *et al.*¹⁰. Both of the *Arctica islandica* fragments analysed on the accelerator yielded radio-

carbon ages >40 kyr (Table 2); the late-glacial control samples had been dated previously by conventional radiocarbon dating (Table 2; ref. 11). Although the Rhubodach site is undated, the fauna and stratigraphy confirm that it is correlative with the shells from Ardyne Point.

The equations describing the isoleucine epimerization reaction^{12,13} require an independent evaluation of the Arrhenius parameters for the taxon under study. If the sample age is known and the D/L ratio can be measured, the average temperature experienced by the shell since deposition can be calculated. The accuracy of such calculations is, however, highly dependent on the derived Arrhenius parameters. A 1% error in the activation energy translates to a ± 3 °C error in calculated temperature. Temperature difference calculations, in contrast, are insensitive to Arrhenius parameters. A 10% error in activation energy results in a discrepancy of less than 1 °C in the calculated temperature difference. The forward rate constant (k) is calculated from

$$k = \frac{\ln\{(1+B)/(1-K'B)\} - C}{(1+K')t} \quad (1)$$

where k is the forward rate constant, B is the D/L ratio, K' is the reciprocal of the equilibrium constant (0.77), C is a constant related to laboratory-induced epimerization (0.025), and t is the interval in years. The relationship defining the dependency of the rate constant, k , on temperature is given by the Arrhenius equation

$$k = S e^{(-E_a/RT)} \quad (2)$$

where S is the intercept value, E_a is the energy of activation, R is the gas constant (1.987) and T is temperature (K). The temperature difference can be calculated from

$$\Delta T = \frac{E_a}{R} \left\{ \frac{1}{A - \ln(k_1)} - \frac{1}{A - \ln(k_2)} \right\} \quad (3)$$

where ΔT is the difference in temperature between two intervals, A is $\ln(S)$, and k_1 and k_2 are the forward rate constants calculated from equation (1) for the two intervals. ΔT can be read directly in degrees Celsius.

We have two sets of samples of known (or minimum) age and D/L ratio (Tables 1 and 2). Using reasonable Arrhenius parameters ($E_a = 28.1$ kcal mol⁻¹; $A = 37.5$) in the equations above, we can compute the difference in temperature between that experienced by the late-glacial samples (essentially the average Holocene temperature) and the period between 11 kyr ago and the age of the King Edward shells. Note that although the late-glacial and King Edward sites lie on opposite sides of Scotland, the present temperature difference is slight (mean annual temperatures are ~8.5 °C around Ardyne and 9.0 °C near King Edward). Because the radiocarbon dates on the King Edward shells are minimum estimates, the computed temperature difference will also be a minimum estimate. In Table 3 we assign various older ages to the King Edward samples to demonstrate the sensitivity of the calculation to sample age. We can constrain the maximum age of the King Edward shells by comparing their D/L ratios with those in last interglacial *A. islandica* from the UK (~0.20). If the King Edward shells are

Table 2 Radiocarbon dates for *Arctica islandica* from localities in Scotland

Locality	Material	Date	Laboratory identification
Ardyne	<i>Arctica islandica</i>	11,160 ± 150	SRR-483
Elf Pit			
Ardyne	<i>Arctica islandica</i>	11,575 ± 100	SRR-481
Shell Pit			
King Edward (AAL 4361)	<i>Arctica islandica</i>	>44,200	AA-1323
King Edward (AAL-4434)	<i>Arctica islandica</i>	>41,500	AA-1324

Table 3 Amino-acid-derived temperature depressions for the Devensian of Scotland

Mean D/L	Known age (kyr)	Assumed age (kyr)	Time range of calculation (kyr BP)	Temperature difference from the Holocene average (°C)
0.055	11			
0.078	>42	42	11-42	-9
0.078	>42	60	11-60	-11
0.078	>42	80	11-80	-13

even 80 kyr old, the time interval between 80 and 120 kyr would have to have been warmer than the Holocene (to explain the rapid epimerization that would be required), a conclusion that is in conflict with other records. Consequently we only consider ages between 40 and 80 kyr as feasible.

The calculated minimum glacial-age temperature reduction across the northern UK is 9°C. The vegetation in Devensian sites in the UK is dominated by treeless taxa associated with periglacial landforms, supporting a considerable temperature reduction. Oceanic surface waters adjacent to the British Isles were at least 10°C colder at the last glacial maximum¹⁴ and the presence of permafrost implies an even greater reduction of the average terrestrial temperature. Because the racemization-derived temperature decrease reflects an integrated temperature for the entire interval between at least 42 kyr and 11 kyr ago, there may have been periods of even more intense cold within this interval. For example, a 10 kyr interval with temperature 20°C lower than the Holocene would alter the temperature depression for the remaining interval by only 2°C (to 7°C). However, because the epimerization reaction responds exponentially to temperature, it is not possible for there to have been any prolonged warm episodes: assuming that the shells are only slightly older than 40 kyr, then 'warm' intervals totalling about 5 kyr duration are the most that could be reasonably accommodated. Even 5 kyr of temperatures similar to the Holocene requires the remainder of the period to have averaged 12°C rather than 9°C lower than the Holocene. A 12°C average temperature depression is plausible, but a greater decrease is unrealistic. This conclusion is compatible with Coope's reconstruction¹⁵ of the Upton Warren complex, suggesting a 2–3 kyr warm excursion just before 40 kyr ago within a period of otherwise greatly depressed summer temperatures. If the shells are considerably older than 40 kyr then the required temperature shifts are even more dramatic.

Isoleucine epimerization ratios in radiocarbon-dated late-glacial and middle Devensian *Arctica islandica* from Scotland place constraints on the temperature over this region between ~40 kyr and 11 kyr BP. The average terrestrial temperature was depressed at least 9°C below the Holocene; some intervals within this period may have been substantially colder, but there could not have been >5 kyr during which temperatures approached those of modern times. The substantial depression of the average temperature in the late middle Devensian and throughout the late Devensian requires that the Atlantic circulation was fundamentally different than now throughout this interval, a conclusion consistent with the findings of the CLIMAP project.

This research was supported by NSF grant AEO-8319032. We thank D. Goter (USA) and G. Sykes (UK) for assistance with the amino acid analyses and L. Toolin for help in the TAMS facility.

Received 10 October 1986; accepted 6 January 1987.

1. Miller, G. H. & Mangerud, J. *Quat. Sci. Rev.* 4, 215–278 (1986).
2. Bowen, D. Q. *et al. Quat. Sci. Rev.* 4, 279–318 (1986).
3. Hearty, P. J., Miller, G. H., Stearns, C. E. & Szabo, B. J. *Bull. geol. Soc. Am.* 97, 850–858 (1986).
4. Wehmiller, J. F. in *Quaternary Dating Methods* (ed. Mahaney, W. C.) 171–193 (Elsevier, Amsterdam, 1984).
5. McCoy, W. D. *Quat. Sci. Rev.* 6 (in the press).
6. Jamieson, T. F. *Q. J. geol. Soc. Lond.* 22, 261–281 (1986).
7. Read, H. H. *Mem. geol. Surv. Scotl.* (1923).
8. Sejrup, H. P. *Abstr. 14th Arctic Workshop, BIO, Halifax, Nova Scotia*, 219–223 (1985).
9. Jull, A. J. T., Donahue, D. J., Hatheway, A. L., Linick, T. W. & Toolin, L. J. in *Proc. 12th Int. Radiocarbon Conf.*, 24–28 June 1985, Trondheim, Norway (eds Stuiver, M. and Kra, R.) 191–197 (*Radiocarbon* 28(2A), 1986).
10. Linick, T. W., Jull, A. J. T., Toolin, L. J. & Donahue, D. J. in *Proc. 12th Int. Radiocarbon Conf.*, 24–28 June 1985, Trondheim, Norway (eds Stuiver, M. and Kra, R.) 522–533 (*Radiocarbon* 28(2A), 1986).
11. Peacock, J. D., Graham, D. K. & Wilkinson, I. P. *Inst. geol. Sci. Rep.* 78/17 (1978).
12. Schroeder, R. A. & Bada, J. L. *Earth Sci. Rev.* 12, 347–391 (1976).
13. Miller, G. H., Sejrup, H. P., Mangerud, J. & Andersen, B. G. *Boreas* 12, 107–124 (1983).
14. CLIMAP Project Members in *Map and Chart Series 36* (A. McIntyre, Leader) (Geological Society of America, 1981).
15. Coope, G. R. in *British Quaternary Studies* (ed. Shotton, F. W.) 55–68 (Clarendon Press, Oxford, 1977).

The ecological niche of an extinct group of mammals, the early Tertiary apatemyids

W. v. Koenigswald* & H.-P. Schierning†

* Hessisches Landesmuseum, Friedensplatz 1, D-6100 Darmstadt, FRG

† Wittland 59 c, D-2000 Hamburg 55, FRG

Apatemyids are highly specialized eutherian mammals from the Paleogene of North America and Europe. They are known almost exclusively from their dentition^{1–3}. Only two skulls have ever been described, and the postcranial skeleton has remained a complete mystery. Reconstructions of apatemyid adaptations for a specific locomotion or diet had to be speculative, even when some similarities to the dentition of the lemuriform primate *Daubentonia* and the marsupial *Dactylopsila* were noticed^{1,4,5}. Now, three nearly complete skeletons are available, and they show an extreme elongation of selected fingers. Thus, apatemyids are a third group of mammals (with *Dactylopsila* and *Daubentonia*⁶) convergently specialized for feeding on wood-boring insects.

The Eocene lake deposits of Messel (near Darmstadt, West Germany), known for the unusually fine preservation of mammalian skeletons^{7,8}, have yielded three skeletons of a small apatemyid in the course of private and scientific excavations.

Because the oil shale deteriorates when it dries, the skeletons have to be transferred to an artificial matrix (epoxy-resin). This means that all elements of the fragile skeleton are visible but cannot be handled independently or studied from all sides. The jaws are normally in occlusion, so that the occlusal surface of the dentition is hidden. The bones are flattened by compaction of the oil shale, but the skeleton as a whole is in a very natural pose, as is often the case in drowned mammals^{9,10} (Fig. 1). All

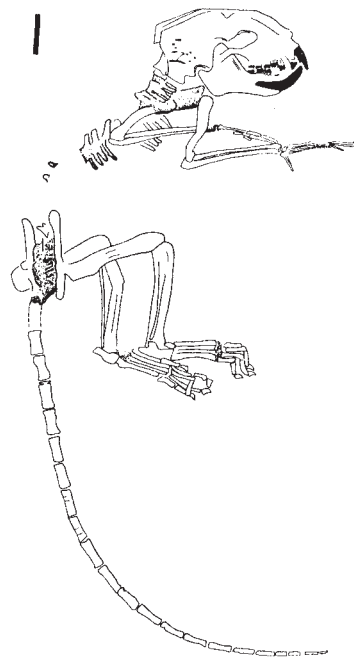
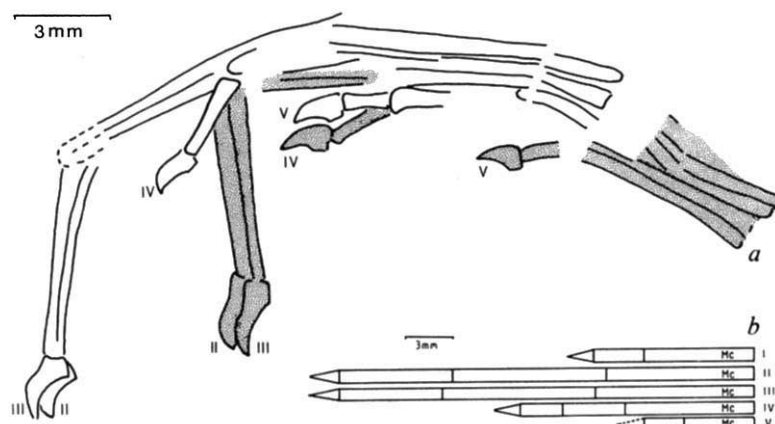


Fig. 1 Skeleton of *Heterohyus* sp. from the middle Eocene oil shale from Messel (near Darmstadt, West Germany). Scale bar, 1 cm. This skeleton (specimen 2, found in 1984) shows the highly specialized hands with elongated fingers, similar to the lemuriform primate *Daubentonia* or the marsupial *Dactylopsila*. The natural pose of the fossil skeleton indicates that this animal was drowned (Landessammlungen für Naturkunde, Karlsruhe, specimen LNK—Me 689).



Fig. 2 Anterior part of crushed skull of *Heterohyus* sp. from Messel (specimen 1, found in 1973), showing the enlarged anterior incisors and the large triangle-shaped second incisors, as well as the relatively small molars (collection of H. P. S.).

Fig. 3 Hands of *Heterohyus* sp. from Messel, drawn from the third specimen. Scale bars, 3 mm in both *a* and *b*. *a*, Right (stippled) and left (white) hand in natural position. The second and third fingers are extremely elongated but are of equal thickness. *b*, Schematic proportions of the various bony elements of the hand (Hessisches Landesmuseum, Darmstadt specimen HLMD—Me 8850).



three individuals are same size and belong to the same small species of the genus *Heterohyus*, but a specific determination cannot be given yet. *Heterohyus* is derived among apatemyids in having strongly curved incisors, which appear to be almost evergrowing (Fig. 2).

The skull of each specimen is about 37–40 mm long and heavily built. Head and body length is about 115 mm and the tail adds another 150 mm. The tail, containing about 22 vertebrae, does not appear to be prehensile. Of special interest are the fore- and hindlimbs. The femur and the tibia are each 28 mm long and strongly built. The first toe is not enlarged and was not opposable, although the spreading of the metatarsals would have made some additional grip possible¹¹. The last phalanges of each toe are formed as high and laterally flattened claws. Even when the ankle joints are in close contact, their position makes it highly probable that the tarsal movement allowed hyperinversion for a safe grip during head-first descent, as in several highly arboreal mammals^{12–14}.

The forelimbs are more gracile than the hindlimbs and are distinctly elongated. The humerus and radius are each 20 mm long, and the ulna is free and slender. The hands are visible in the second and third specimen and reveal an unusual specialization of the fingers. The first and fifth fingers are short, the fourth is of intermediate length, and the second and third fingers are greatly elongated. Together with the metacarpals, their length, measured from the third specimen during preparation, is about 26.5 mm. All the fingers bear claws.

Enlarged procumbent incisors, together with an elongation

of selected fingers in the hand, as shown in *Heterohyus* sp., are a very distinct combination of characters found in two extant mammals, the marsupial *Dactylopsila trivirgata* and the lemuriform primate *Daubentonia madagascariensis*. Both species have developed these structures to feed on wood-boring insects. The enlarged incisors are used to tear off bark and to open living chambers of beetle larvae. The elongated fingers are used to extract the larvae from their burrows or from narrow fissures, as well as to anchor the animal while it handles its prey¹⁵.

These anatomical characters of *Dactylopsila* and *Daubentonia* are strongly related to the highly specialized mode of feeding and are shared with *Heterohyus*. These characters arose through convergent evolution in the three genera and therefore it can be deduced that *Heterohyus* sp. was probably feeding on wood-boring insects in a similar way.

It is characteristic of convergence that, despite important similarities, certain differences can be recognized. In this case, digital elongation affects different fingers. In *Dactylopsila* it is the fourth finger, whereas in *Daubentonia* it is the third which is elongated. This extremely thin finger is supported by the much thicker fourth finger of equal length in *Daubentonia*. In *Heterohyus* sp. both the second and the third fingers are elongated, and they are of equal thickness. Because a characteristic dentition was developed throughout the apatemyids, it seems plausible that they were all specialized for essentially the same ecological niche, even assuming different degrees of adaptation.

Cartmill¹⁶ demonstrated that New Guinea (with *Dactylopsila*)

and Madagascar (with *Daubentonia*) are the only extensive areas that are not colonized by woodpeckers, so that this niche is available for specialized mammals. By analogy, it might be concluded that apatemyids and woodpeckers exclude each other, and that woodpeckers therefore were not present during the Paleogene in Europe and North America.

Although fossil bones of woodpeckers date back only to the Miocene, a probable woodpecker's hole preserved in silicified wood from the Eocene of Arizona¹⁷ was described recently. Therefore, it can not be excluded that some primitive woodpeckers lived sympatrically with apatemyids in the Eocene.

The authors thank Dr J. Hooker, London and Dr W. P. Luckett, San Juan, for reading the manuscript and for their valued comments.

Received 11 August 1986; accepted 16 February 1987.

- West, R. M. *Spec. Publ. Mus. Texas Tech. Univ.* **3**, 1-42 (1973).
- Russell, D. E. *et al. Bull. Mus. natn. Hist. nat.* **1C**, 203-243 (1979).
- Gingerich, P. D. & Rose, K. D. *Contr. Mus. Paleont. Univ. Mich.* **26**, 49-55 (1982).
- Stehlin, H. G. *Abh. schweiz. paläont. Ges.* **41/7**, 1229-1552 (1916).
- McKenna, M. C. *Am. Mus. Novit.* **2160**, 1-39 (1963).
- Starck, D. *Vergleichende Anatomie der Wirbeltiere* Vol. 2 (Springer, Berlin, 1979).
- Koenigswald, W. v. & Michaelis, W. *Geol. Jb. Hessen* **112**, 5-26 (1984).
- Franzen, J. L. *Phil. Trans. R. Soc. B311*, 181-186 (1985).
- Weigelt, J. *Reizente Wirbeltierleichen* (Max Weg, Leipzig, 1927).
- Koenigswald, W. v. *Neues. Jb. Geol. Paläont. Abh.* **167**, 1-39 (1983).
- Napier, J. R. & Napier, P. H. *A Handbook of Living Primates* (Academic, London, 1967).
- Jenkins, F. A. & Krause, D. W. *Science* **220**, 712-715 (1983).
- Jenkins, F. A. & McClearn, D. J. *Morphology* **182**, 197-219 (1984).
- Szalay, F. S. in *Evolutionary Relationships Among Rodents* (eds Luckett, W. P. & Hartenberger, J. L.) 83-132 (Plenum, New York, 1985).
- Moeller, H. *Publ. n. wiss. Film. Sektion Biol.* **9/2**, 129-138 (1976).
- Cartmill, M. in *Prosimian Biology* (eds Martin, R. D. & Doyle, G. A.) 655-669 (Duckworth, London, 1974).
- Buchholz, H. *Verh. naturw. Ver. Hamb.*, NF **28**, 5-25 (1986).

Enigmatic changes in the hydromedusan fauna of the northern Adriatic Sea

Adam Benović*, Dubravko Justić† & Ankica Bender*

* Biological Institute, PO Box 39, 50000 Dubrovnik, Yugoslavia

† Department of Zoology, University of Zagreb, Rooseveltov trg 6, 41000 Zagreb, Yugoslavia

In many coastal and estuarine waters^{1,2}, sensitive bottom-dwelling animals are threatened by depletion of the oxygen concentration in the near-bottom layer³⁻⁵, as a result of eutrophic conditions inshore. Persistent or occasional periods of low oxygen can reduce the number of benthic species, as has been reported for Chesapeake Bay⁶. Although it has been hypothesized^{6,7} that low oxygen concentration in the near-bottom layer extirpates those plankton species that have a bottom phase in their life cycles (meroplankton), there has been no direct evidence for this. We present here a comprehensive set of data which clearly demonstrates a strong correlation between oxygen depletion near the bottom and the disappearance of meroplanktonic hydromedusan species from the fauna of the northern Adriatic Sea. We believe that we have identified indicator species of oxygen deficiencies in marine ecosystems.

The northern Adriatic Sea has been studied over a long period^{8,9}, and its plankton fauna is regarded as typically neritic¹⁰⁻¹⁴. Occasional phytoplankton blooms in this sea have long been known^{15,16}, but during the past 20 years the northern Adriatic has become increasingly eutrophic¹⁷⁻¹⁹. Although changes in the phytoplankton community have been related to eutrophication and pollution¹⁸⁻²⁰, there have been no previous reports of a possible relation between oxygen depletion in the near bottom layer and changes in the zooplankton composition.

Plots of temperature and salinity during the summer from 1911 to 1982 (Fig. 1) show no apparent increasing or decreasing

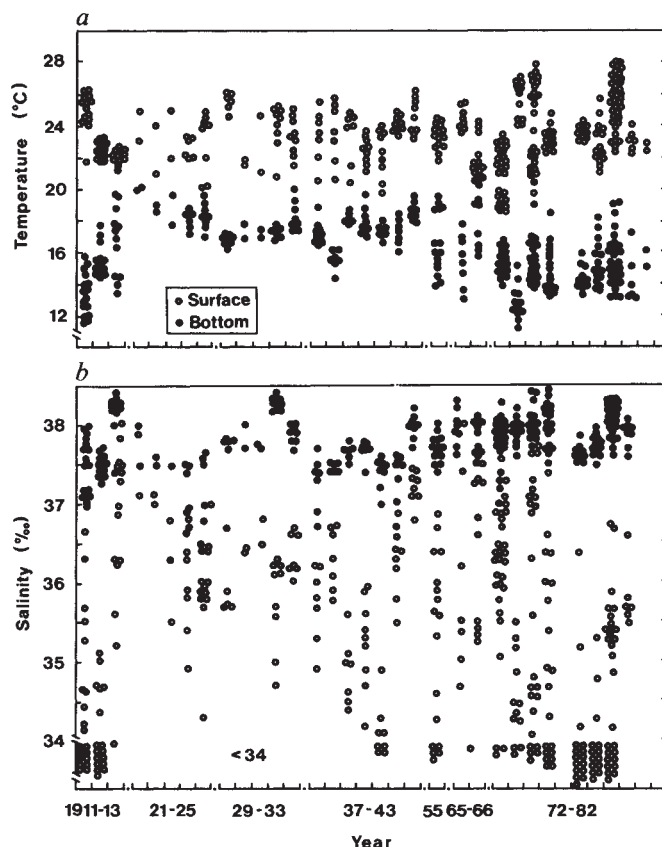


Fig. 1 Temperature (a) and salinity (b) in the northern Adriatic Sea during August and September, from 1911 to 1982. Surface (○) and bottom (●) samples were collected at 0.5 m below the surface and 2.0 m above the bottom, respectively. The method was constant throughout the period; for details of the sampling and chemical analyses as well as a listing of the data see ref 5. Additional unpublished data were provided by Institutes acknowledged in the text.

trend, either at the surface or near the bottom, although the data are highly varied during this period. In contrast, the plot of oxygen saturation in the same summers (Fig. 2a) demonstrates a trend towards lower values near the bottom, first visible in the 1955 data. After 1965, the tendency towards oxygen depletion of the near-bottom water and hypersaturation near the surface became increasingly more apparent. The change has been substantial. Calculated from D.J., T. Legović and L. Rotini-Sandrini, (manuscript in preparation), the oxygen saturation near the surface increased by an average of 0.25% per year and decreased near the bottom at an average of 0.5% per year. Further, the O₂ concentration of the near-bottom layers approached the critical level at which the presence of H₂S could be expected^{3,5}, as already reported for the Bay of Trieste²¹. This conclusion is reinforced by the recent occurrence of local benthic mortalities in the Bay of Trieste^{4,21,22} and near the western middle-Adriatic coastline^{20,23,24}. Increasing oxygen saturation near the surface coincides with the recent increase in frequency of phytoplankton blooms²⁵ and can be attributed to increasing eutrophication (ref. 17; D.J. *et al.*, manuscript in preparation). Further, some recent provisional oxygen data from near the Po estuary show the lowest values to date for winter²⁶.

The hydromedusan fauna of the northern Adriatic Sea, like that of other neritic ecosystems, usually comprises a large number of Anthomedusae and Leptomedusae. These species have alternating pelagic (hydromedusa) and bottom-dwelling (hydroid) phases in their life cycles (which are thus metagenetic). The other hydromedusan orders (Trachymedusae and

Table 1 Historical changes in the hydromedusa fauna of the northern Adriatic Sea

Anthomedusae	1910	1965	1972	1974/1976	1984/1985	Leptomedusae	1910	1965	1972	1974/1976	1984/1985
1. <i>Dicodinium adriaticum</i>	+	+				29. <i>Orchistomella graeffei</i>	+				
2. <i>Dipurena halterata</i>	+	+				30. <i>Laodicea undulata</i>	+				
3. <i>Sarsia gemmifera</i>	+	+	+	+		31. <i>Obelia</i> spp.	+	+	+	+	+
4. <i>Stauridiosarsia producta</i>	+	+				32. <i>Clytia hemisphaerica</i>	+	+	+	+	+
5. <i>Eucodinium browni</i>	+	+				33. <i>Eucope picta</i>	+				
6. <i>Ectopleura dumortieri</i>	+	+	+	+		34. <i>Eucheilota maasi</i>	+				
7. <i>Euphysa aurata</i>		+	+	+		35. <i>Eirene viridula</i>	+				
8. <i>Rhabdoon singulare</i>			+	+		36. <i>Helgicirrho schultzei</i>	+	+		+	
9. <i>Corymorpha nutans</i>	+	+	+	+		37. <i>Eutima gegenbauri</i>	+	+		+	
10. <i>Zanclea costata</i>	+	+				38. <i>Eutima gracilis</i>			+		+
11. <i>Cladonema radiatum</i>	+					39. <i>Eutonina scintillans</i>	+				
12. <i>Eleutheria dichotoma</i>	+					40. <i>Tima luculana</i>	+				
13. <i>Cytaeis tetrastyla</i>		+				41. <i>Aequorea aequorea</i>	+				
14. <i>Turitopsis nutricula</i>	+										
15. <i>Podocoryne carnea</i>	+	+				Trachymedusae					
16. <i>Podocoryne areolata</i>	+					44. <i>Geryonia proboscydalis</i>	+				
17. <i>Podocoryne minima</i>		+	+	+	+	45. <i>Liriope tetraphylla</i>	+	+	+	+	+
18. <i>Podocoryne minuta</i>	+	+	+	+	+	46. <i>Aglaura hemistoma</i>	+	+	+	+	+
19. <i>Rathkea octopunctata</i>	+					47. <i>Persa incolorata</i>		+		+	+
20. <i>Bougainvillia ramosa</i>	+	+	+			48. <i>Rhopalonema velatum</i>	+	+	+		+
21. <i>Lizzia octostyla</i>	+										
22. <i>Lizzia blondina</i>	+	+	+			Narcomedusae					
23. <i>Thamnostoma dibalia</i>	+	+				49. <i>Solmundella bitentaculata</i>	+	+		+	+
24. <i>Merga tergestina</i>	+	+				50. <i>Solmaris leucostyla</i>	+	+	+	+	+
25. <i>Amphinema dinema</i>	+					51. <i>Solmaris vanhoeffeni</i>	+	+	+		
26. <i>Leuckartiara octona</i>			+	+							
27. <i>Pandea</i> sp.	+					Total species	42	29	18	18	11
28. <i>Neoturris pileata</i>	+	+		+							

Historical data on hydromedusae¹⁰ used as qualitative reference include the most comprehensive set of 1909–1911 data³⁵. Hydromedusae have been collected with the same methods since 1965^{10,27}; the quality of sampling efforts before that date is not comparable. The decline in number of species is valid for the entire period from 1965 onwards, and strictly speaking the set of data from 1909–1911 serve only as a reference for number of species. For 1910 samples, see ref. 35; for 1965, ref. 10; and for 1974/1976 and 1984/1985, see ref. 27.

Table 2 Seasonal occurrence of hydromedusa species in the Northern Adriatic Sea during the period 1972–1985

	1972								1974	1975	1976	1976	1984	1985
	III	V	VI	VII	VIII	IX	X	IX/X	IV	II	VI	XII	VIII	
Anthomedusae														
3. <i>Sarsia gemmifera</i>	+	+	+	+	+	+	+		+		+			
6. <i>Ectopleura dumortieri</i>				+				+						
7. <i>Euphysa aurata</i>				+				+	+					
8. <i>Rhabdoon singulare</i>	+	+	+	+	+	+	+	+						
9. <i>Corymorpha nutans</i>	+	+							+		+			
17. <i>Podocoryne minima</i>	+		+	+	+	+	+	+	+		+	+	+	
18. <i>Podocoryne minuta</i>	+	+	+	+	+			+	+	+	+	+		
20. <i>Bougainvillia ramosa</i>	+			+	+									
22. <i>Lizzia blondina</i>				+										
26. <i>Leuckartiara octona</i>	+				+	+				+				
28. <i>Neoturris pileata</i>									+					
Leptomedusae														
31. <i>Obelia</i> spp.	+	+		+	+	+	+	+	+	+	+	+	+	
32. <i>Clytia hemisphaerica</i>				+	+			+	+	+	+	+	+	
36. <i>Helgicirrho schulzei</i>								+						
37. <i>Eutima gegenbauri</i>										+				
38. <i>Eutima gracilis</i>					+								+	
Trachymedusae														
45. <i>Liriope tetraphylla</i>				+	+		+	+	+	+	+	+	+	
46. <i>Aglaura hemistoma</i>	+		+		+	+	+		+	+	+	+	+	
47. <i>Persa incolorata</i>									+	+		+		
48. <i>Rhopalonema velatum</i>							+					+		
Narcomedusae														
49. <i>Solmundella bitentaculata</i>										+		+		
50. <i>Solmaris leucostyla</i>								+	+				+	
52. <i>Solmaris</i> spp.*	+	+	+	+	+	+	+							
53. <i>Hydromedusae</i> juv.	+	+		+	+									
Total species	11	7	6	12	14	7	8	9	12	10	7	9	6	

* Comprises *S. leucostyla* and *S. vanhoeffeni* which, because of poor preservation, were not separable. Roman numerals represent month of sampling. See Fig. 1 and Table 1 for additional explanations and data sources.

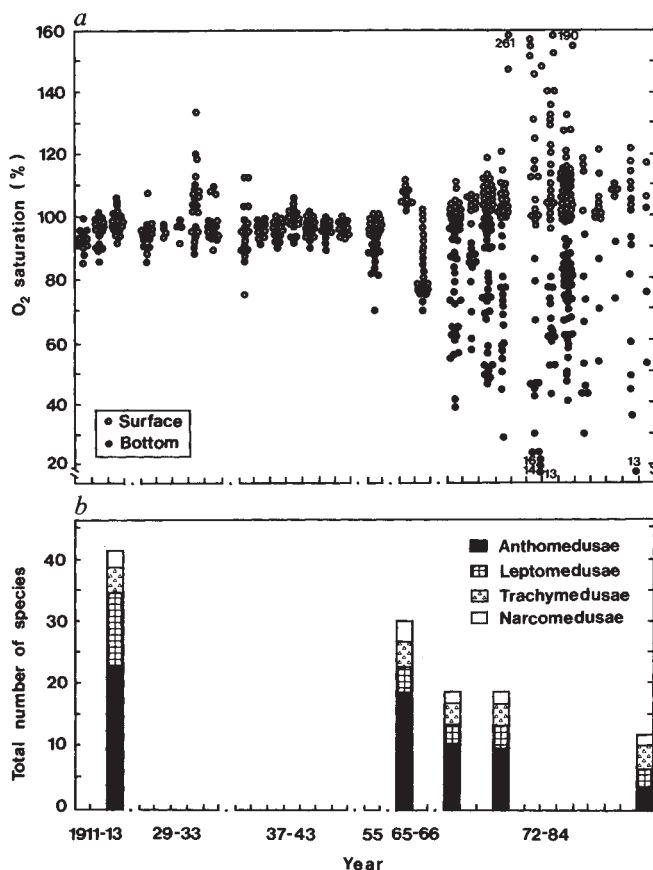


Fig. 2 a, Surface (○) and bottom (●) oxygen saturation during summer, and b, total number of hydromedusan species, in the northern Adriatic Sea from 1911 to 1984. Each 10% of saturation is approximately equal to $0.55 \text{ cm}^3 \text{O}_2 \text{ dm}^{-3}$. See Fig. 1 and Table 1 for additional explanation and data sources.

Narcomedusae) do not need the bottom for reproduction, because there is no hydroid stage in their life history. Assessment of the changing composition of the hydromedusan fauna over the period 1911–1985 reveals a substantial decline in the total number of species: 22 anthomedusan and 9 leptomedusan species have completely disappeared from the fauna of the northern Adriatic Sea (Fig. 2b, Table 1). Although the decrease in leptomedusan species may coincide with the oxygen data separation after 1955 (Fig. 2a), the greatest decrease in Anthomedusae correlates with the trend of near-bottom oxygen depletion after 1965. The greatest impoverishment in meta-genetic species numbers was recorded in the summer of 1985 (Table 2). In contrast, Trachymedusae and Narcomedusae remained nearly at the same number of species throughout the investigated period (Fig. 2b, Table 1). Their abundance is related to water-mass movement in the middle and southern Adriatic²⁷, and is hence not affected by near-bottom oxygen depletion in the north.

We have no relevant data for hydroids and hydromedusae, but we hypothesize that these animals are extremely sensitive to oxygen depletions whether it is caused by natural phytoplankton blooms or by pollution^{2,28}. The correlation between the total number of anthomedusan and leptomedusan species and oxygen saturation near the bottom (Fig. 2) supports this hypothesis. The low oxygen concentrations near the bottom that have been reported in recent years and which persist for 2–3 months during the summer (ref. 21; D.J. *et al.*, manuscript in preparation), may affect proliferation of the hydropolyps. The most recent impoverishment of oxygen was reported in November 1985²⁶. Other factors that may have contributed to that decrease in O_2 tension are the pronounced sinking of organic matter from the pelagic

strata²⁹ (known to contribute to oxygen depletion³⁰) and silting of the bottom substrata²⁸ from the Po river sediment discharge²⁰. However, there is no trend of decrease in surface salinity (Fig. 1b). Although the quality of early hydromedusan sampling can be questioned, the sampling effort was considerably increased after 1972, so the loss of species cannot be attributed to inadequate sampling. Lastly, the most recent decrease in the hydromedusan fauna may be in part influenced by an increase in density of the scyphozoan *Pelagia noctiluca*³¹, which preys on hydromedusae³². Large swarms of *Pelagia noctiluca* were first recorded in the northern Adriatic Sea in 1977³¹; before 1977 it was rare in this region and therefore cannot be responsible for the decrease in the number of hydromedusan species before 1977.

It is possible that the numbers of anthomedusan and leptomedusan species may have been reduced by an interplay of several factors, but it is hard to believe that something other than oxygen was the dominant one.

In collecting hydrographic data by sampling, sudden and brief near-bottom anoxic events, during which sensitive animals might be killed, may be missed. Nevertheless, the near-total disappearance of whole orders of animals is probably a sign of changes in the ecosystem. Progressive reduction in oxygen content could well cause mass faunal or zooplankton mortalities^{1–3,23}. Relating oxygen depletion trends in the near-bottom regions to hydrographic currents in the Adriatic Sea^{33,34}, we predict mass mortalities of hydromedusae and perhaps of many other animal groups in the offshore northern Adriatic Sea on a large scale in the near future.

We thank Dr P. F. S. Cornelius, British Museum (Natural History), London and Dr P. H. Wiebe, Woods Hole Oceanographic Institution for their criticism and help with the English manuscript. We also thank Dr D. Degobbi, of the Ruder Bošković Institute, Centre for Marine Research, Rovinj, and Dr A. Malej and Dr N. Fanuko, Institute of Biology, University of Ljubljana MRIC, Piran, Yugoslavia, for recent hydrographic data and plankton samples from 1972, respectively. This work was supported by the Self-Management Community of Interest for Scientific Research of SR Croatia, Yugoslavia, and the United Nations Environmental Programme, Mediterranean Unit, Athens, Greece.

Received 4 August 1986; accepted 9 February 1987.

- Brongersma-Sanders, M. *Mem. geol. Soc. Am.* **67**, 941–1010 (1957).
- Falkowski, P. G., Hopkins, T. S. & Walsh, J. J. *J. mar. Res.* **38**, 479–506 (1980).
- Theede, H., Ponat, A., Hiroki, K. & Schlieper, C. *Mar. Biol. Berlin* **2**, 325–337 (1969).
- Ott, J. & Fedra, K. *Helgoländer wiss. Meeresunters.* **30**, 485–494 (1977).
- Richards, F. A. in *Chemical Oceanography* Vol. 1 (eds Riley, J. P. & Skirrow, G.) 611–645 (Academic, London, 1965).
- Officer, C. B. *et al. Science* **223**, 22–27 (1984).
- Davis, C. O., Hollibaugh, J. T., Seibert, D. L. R., Thomas, W. H. & Harrison, P. J. *J. Phycol.* **16**, 296–302 (1980).
- Buljan, M. & Zore-Armanda, M. *Oceanogr. mar. Biol. A. Rev.* **14**, 11–98 (1976).
- Ghirardelli, E. *Thalassia jugosl.* **19**, 151–172 (1983).
- Benović, A. *Boll. Pesca Piscic. Idrobiol.* **28**, 59–70 (1973).
- Specchi, M. *Boll. Pesca Piscic. Idrobiol.* **25**, 95–110 (1970).
- Fenaux, R. *Mar. Biol., Berlin* **16**, 310–319 (1972).
- Hure, J., Ianora, A. & Scotto di Carlo, B. *J. Plankton Res.* **2**, 295–316 (1980).
- Benović, A., Fonda-Umani, S., Malej, A. & Specchi, M. *Mar. Biol., Berlin* **79**, 209–218 (1984).
- Cori, C. J. *Arch. Hydrobiol.* **1**, 38–42 (1906).
- Steuer, A., Teubner, B. G. *Planktonkunde* 1–382 (Leipzig, Berlin, 1910).
- Degobbi, D., Smolaka, N., Pojed, I., Skrivanić, A. & Precali, R. *Mar. Poll. Bull.* **10**, 298–301 (1979).
- Gilmartin, M. & Revelante, N. *Mar. Ecol. Prog. Ser.* **1**, 169–180 (1980).
- Revelante, N., Gilmartin, M. & Smolaka, N. *J. Plankton Res.* **7**, 461–471 (1985).
- Chiaudani, G., Gaggino, G. F. & Vighi, M. *Thalassia jugosl.* **19**, 77–86 (1983).
- Faganeli, J. *et al. Mar. Pollut. Bull.* **16**, 75–78 (1985).
- Stachowitsch, M. *Publ. Staz. zool. Napoli, I: mar. Ecol.* **5**, 243–264 (1984).
- Frogia, C. *Giorn. Bot. Ital.* **104**, 389–396 (1970).
- Piccinetti, C. & Manfrin, G. *Note Lab. Biol. mar. Fano* **3**, 73–92 (1969).
- Fonda-Umani, S. *Oebalia* **11**, 141–147 (1985).
- Tornati, G. *Atti VIII Conf. Delle Città Adriatiche Italiane e Jugoslave* 1–11 (Budva, Jugoslavia, 3–4 April 1985).
- Benović, A. & Bender, A. in *Modern Trends in the Systematics, Ecology and Evolution of Hydroids and Hydromedusae* (eds Bouillon, J., Boero, F., Cicogna, F. & Cornelius, P. F. S.) (Oxford University Press, in the press).
- Boero, F. *Publ. Staz. Zool. Napoli, I: mar. Ecol.* **5**, 93–118 (1984).
- Staršćin, N. *et al. V^{te} Journées Etud. Pollutions, C.I.E.S.M., Cagliari* 1011–1012 (1980).
- Oviatt, C. A., Keller, A. A., Sampou, P. A. & Beatty, L. L. *Mar. Ecol. Prog. Ser.* **28**, 69–80 (1986).
- Vučetić, T. & Rottini-Sandrini, L. *Rapp. Comm. int. Mer. Mediterr.* **28**, 131–132 (1983).

32. Russell, F. S. *The Medusae of the British Isles, II* (Cambridge University Press, 1970).
33. Zore-Armanda, M. *Stud. Rev. gen. Fish. Coun. Mediterr.* **34**, 1-48 (1968).
34. Malanotte-Rizzoli, P. & Bergamasco, A. *J. phys. Oceanogr.* **13**, 1105-1130 (1983).
35. Neppi, V. & Stiasny, G. *Arb. zool. Inst. Univ. Wien* **20**, 23-92 (1913).
36. Bruckner, E. *Berichte über die Terminfahrten 1911-1914, SMS Najade, 2-12 and RN Cyclope 1-3* (Holzhausen, Vienna, 1914).
37. Picotti, M. & Vatova, A. *Thalassia* **5**, 1-157 (1942).
38. Vatova, A. *Nova Thalassia* **1**, 1-63 (1948).
39. Picotti, M. *Archo. Oceanogr. Limnol.* **11**, 371-420 (1960).
40. Trotti, L. *Crociere - Mare Adriatico, 1965-1966, dati oceanografici* (Ist. Sper. Talass., Trieste, 1969).
41. Franco, P. *Archo. Oceanogr. Limnol.* **16**, 1-93 (1970).
42. Franco, P. *Archo. Oceanogr. Limnol.* **17**, 1-97 (1972).
43. Gilmartin, M., Kveder, S., Degobbi, D., Revelante, N. & Smolaka, N. *Thalassia Jugosl.* **8**, 149-213 (1972).
44. Gilmartin, M., Kveder, S., Degobbi, D., Revelante, N. & Smolaka, N. *Thalassia Jugosl.* **8**, 331-438 (1972).
45. Cescon, B. & Scarzatto, P. *Boll. Geof. Teor. ed. Appl.* **21**, 13-37 (1979).
46. Franco, P. *Archo. Oceanogr. Limnol.* **20**, 1-207 (1982).
47. Andrija Mohorovičić 1974-1976. *Reports and Results of the Oceanographic Investigations in the Adriatic Sea*. 228 (Hydrographic Inst. of the Yugoslav Navy, Split, 1982).

cGMP-dependent cation channel of retinal rod outer segments

Diane Matesic & Paul A. Liebman

Departments of Biochemistry and Biophysics and Anatomy, University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania 19104, USA

Light-modulated cytoplasmic cGMP simultaneously controls plasma membrane Na^+ conductance in visual excitation and Ca^{2+} entry into rods by direct interaction with the cation channel¹⁻³. Cytoplasmic Ca^{2+} in turn may set operating points and contribute to the dynamics of several enzymes that regulate cGMP levels in the dark⁴⁻⁶, recovery from excitation and receptor adaptation or down regulation⁷. Similar channels may couple electrical activity to internal nucleotide metabolism in other tissues. We here report the identification, partial purification and behaviour after reconstitution of a protein of relative molecular mass 39,000 (M_r 39K) present in both disk and plasma membranes from bovine rod outer segments that mediates these cGMP-dependent cation fluxes. Its cGMP agonist specificity, kinetic cooperativity, ionic selectivity, membrane density and other features closely match the properties of the visual cGMP-dependent conductance inferred from electrophysiological measurements^{1,2,8,9}.

Rod disk membranes (RDM) prepared from light-adapted bovine retinas¹⁰ were stripped of peripheral proteins (giving SRDM) by washing in hypotonic pH 7.5 Tris buffer containing 50 μM GTP. Monitoring of cGMP-dependent cation fluxes was as described in Fig. 1. Fluxes were in the direction of decreasing cation concentration and decreased with approach to concentration equilibrium. Neither cGMP nor 8-bromo-cGMP trapped in the internal compartment of disks could mediate a cation flux increase. Thus, agonist binding sites are confined to the cytoplasmic face. Influx measurements on four separate membrane preparations gave a cGMP $K_{1/2}$ of $62.5 \pm 15.1 \mu\text{M}$, independent of Ca^{2+} from 50-1,000 μM . The $K_{1/2}$ determinations of cGMP analogues (Fig. 1 legend) reveal the high specificity of the channel for 3', 5'-cGMP and 8-substituted derivatives. Initial rates increased with the external Ca^{2+} concentration with $K_{1/2}$ of 110 μM and maximum Ca^{2+} flux of up to 8.0×10^{-7} mol Ca^{2+} per mg protein per min at 22 °C. Maximum flux at 37 °C was 2-3 times higher. Flux dependence on cGMP concentration (Fig. 1) showed agonist cooperativity, with a Hill coefficient of 1.7 for both SRDM and reconstituted 39K protein vesicles prepared as described below. In both preparations, 2-5 μM 1-*cis*-diltiazem¹¹ inhibited the Ca^{2+} flux by 50%. We also detected cGMP-dependent Mg^{2+} (ref. 12) and Na^+ (ref. 13) fluxes in SRDM. Maximum Ca^{2+} flux in toad (*Bufo marinus*) was comparable to that in bovine SRDM, whereas the cGMP

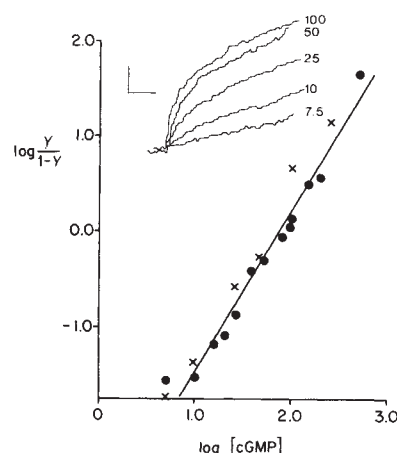


Fig. 1 Hill plot of initial Ca^{2+} flux as a function of $[\text{cGMP}]$ for SRDM ($\bullet$) and for reconstituted 39K protein vesicles (recons $\times$); $Y = (\text{initial flux at the given cGMP concentration})/(\text{maximal flux determined from the intercept of a double reciprocal plot of the data})$. Inset shows typical absorbance traces for recons at the cGMP concentrations shown (in μM); scale bar 1×10^{-3} OD units vertically and 10 s horizontally.

Methods. Vesicles (SRDM or recons) were loaded with 20 mM arsenazo-III (Az III) (ref. 24) in buffer A (100 mM KCl, 20 mM MOPS, 0.5 mM MgCl_2 , pH 7.0) by sonication to trap the dye, followed by washing in buffer A to remove external dye. Vesicles were suspended in buffer A at a protein concentration of 0.15 mg ml^{-1} for SRDM or 0.23 mg ml^{-1} recons. Optical transmission changes due to Ca^{2+} influx were monitored in dual wavelength mode (652-700 nm) at 23 °C. Ion and nucleotide solutions prepared in buffer A were injected (4-10 μl) into 0.4 ml suspensions and rapidly stirred to initiate determinations. The extinction coefficient (E) for Az III in buffer A was determined according to the method of Kendrick¹². Flux densities were calculated from initial slopes using the equation: flux density = $(\text{OD} \times \text{cuvette volume})/E \times t \times [\text{protein}]$; t , time. External Ca^{2+} concentrations used were 100 μM CaCl_2 for SRDM and 1 mM for recons. Initial rate dependence on $[\text{cGMP}]$ was similar for SRDM assayed at 1 mM CaCl_2 . Results are averaged from 2 or 3 determinations at each cGMP concentration. The slope (n) of the apparent best linear fit is 1.74. Values of n obtained from other SRDM and recon preparations ranged from 1.60 to 1.95. The $K_{1/2}$ taken from the illustrated plot is 56 μM cGMP. Estimates of $K_{1/2}$ for other nucleotides determined by Ca^{2+} flux measurements using SRDM are: 8-bromo-cGMP, 15-20 μM ; 8-bromo-cIMP, 200-250 μM ; cIMP, 400-500 μM ; 8-azido-cAMP, 600-800 μM . The $K_{1/2}$ of cAMP was estimated to be 2.0-5.0 mM. On some SRDM preparations, however, cAMP was ineffective at concentrations as high as 2.0 mM. Further, GMP, GTP, GDP, ATP, *N*,*O*-dibutryl-cGMP, *N*-monobutryl-cGMP, *O*-monobutryl-cGMP and 2',3'-cGMP were ineffective at 1.0 mM. In addition, 8-azido-cGMP was reported to have a $K_{1/2}$ of 19 μM (ref. 25). It was determined that absorption changes result from Ca^{2+} -Az III complexing, rather than from pH effects on dye absorption. The cGMP-dependent Ca^{2+} flux measurements were corroborated using ETH-1001-embedded PVC Ca^{2+} -selective electrodes²⁶.

$K_{1/2}$ (40 μM) and 8-bromo-cGMP $K_{1/2}$ (8-10 μM) were somewhat lower.

A membrane component responsible for the cGMP-induced cation flux was identified using the photoaffinity probe, 8-azido-[^{32}P]cGMP, which specifically labelled a protein migrating between rhodopsin and G protein (GTP-binding protein) in our SDS gel system (Fig. 2, lanes *a* and *b*). In most other SDS gel systems, this ~39K protein co-migrated with rhodopsin. Use of 8-azido-[^{32}P]cAMP also labelled a 39K protein in SRDM and in functionally competent reconstituted protein fractions, but much longer autoradiograph exposure times were required to detect binding, because of lower binding affinity. Binding of the photoaffinity reagents could be reduced by addition of unlabelled cGMP before irradiation (50 μM cGMP reduced labelling

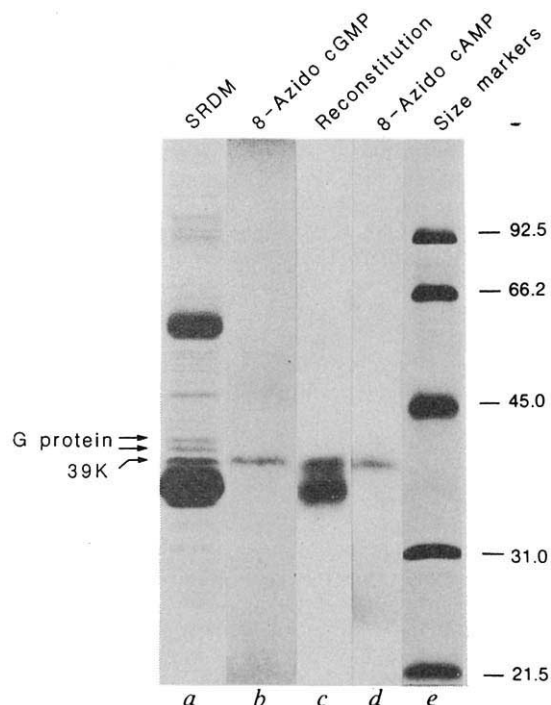


Fig. 2 Coomassie-stained SDS-polyacrylamide gel of 8-azido-[32 P]cGMP-labelled SRDM (a) and corresponding autoradiogram (b) exposed for 32 h. c, d, PC reconstitutions containing 39K protein (c) and autoradiogram of 8-azido-[32 P]cAMP-labelled PC reconstitutions exposed for 210 h (d). e, Size markers, M_r in thousands. The second most prominent band in lane a is rhodopsin dimer with an apparent M_r of 62–63K.

Methods. The SRDM or reconstituted protein fractions were suspended in buffer A at a protein concentration of 0.25–0.90 mg ml $^{-1}$ in 12 $\times$ 75 mm glass tubes. Suspensions (100 μ l) were preincubated for 3 min at 15 $^{\circ}$ C with 10 μ Ci (3.3 μ M) of 8-azido-[32 P]cGMP (30 Ci mmol $^{-1}$, ICN Radiochemicals) or 20 μ Ci (3.0 μ M) of 8-azido-[32 P]cAMP (67 Ci mmol $^{-1}$, ICN). Samples were irradiated with a Hg lamp for 0.5–3.0 min, 2 cm from the top of the tubes, with constant stirring. Reactions were quenched by addition of an equal volume of buffer O (16% glycerol, 0.1 M Tris, 3.5% SDS, 0.25 M DTT, 0.1 g bromothymol blue, pH 6.8). Protein concentrations were determined according to Bradford²⁷ or Sedmark²⁸ except that 0.7% octyl glucoside was added to blanks, bovine serum albumin (BSA) standards, and samples before addition of the Coomassie reagent. SDS-polyacrylamide slab gels were prepared, run and stained as in Fig. 4. Stained gels were either dried or sealed in ziplock plastic bags for autoradiography using Kodak XAR-5 film.

of 8-azido-cGMP twofold in the experiment shown in Fig. 2, lane b). Longer irradiation times resulted in labelling of other SRDM proteins. These other proteins were not detectable in our most highly purified protein fractions that mediated cGMP-dependent Ca $^{2+}$ flux.

The 39K protein could not be washed from SRDM with hypotonic buffers, high salt or mild detergent like 0.1% Nonidet P-40) treatment. During reconstitution, it was partitioned preferentially into the lipid/membrane phase (>90%) and is thus probably embedded in the native membrane. Various ionic and nonionic detergents were screened for stabilization of the cGMP-dependent cation flux on solubilization and reconstitution into phosphatidyl choline (PC) vesicles. We obtained best activity recovery with sodium cholate. The purification method described in Fig. 3 legend yielded a protein fraction containing 39K protein and rhodopsin in a ratio ranging from 1:1 (Fig. 3, lane b) to 1:3 (Fig. 2, lane c) 39K protein/rhodopsin Silver staining¹⁴ to reveal trace contaminants of one such preparation on SDS gels (Fig. 3, lane e) showed no evidence of components of higher M_r . This fraction was able to mediate cGMP-depen-

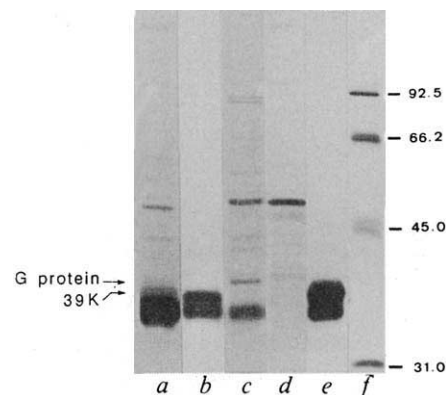


Fig. 3 Comparison of cGMP-dependent Ca $^{2+}$ flux activity and 39K protein enrichment. SDS polyacrylamide gels were run as described for Fig. 4. Maximum Ca $^{2+}$ flux, expressed in mol Ca $^{2+}$ per mg protein per min ($n=1$) were as follows (values in parentheses are normalized for initial SRDM activity and 50% accessible cGMP binding sites in recons ($n=4$) for relative activity comparisons): a, 1.24 (1.00); b, 6.84 (9.24); c, 0.16 (0.11); d–f, not applicable. Extracts in lanes a–e are as follows: a, SRDM used in cholate solubilization; b, 39K protein prep (15 μ g) containing rhodopsin purified by the Sephadex 4B-CL method described below. c, Sephadex 4B-CL column fraction enriched in non-39K protein components. d, Sephadex 4B-CL/Con A Sephadex column fraction containing trace amounts of or no, 39K protein and lacking rhodopsin. e, Silver stain of best 39K-protein-enriched fraction (8 μ g) used in Ca $^{2+}$ flux measurements (included in $n=4$ normalized data). f, Size markers, M_r in thousands.

Methods. SRDM (2.5 mg ml $^{-1}$) were solubilized in 2.0–2.5% sodium cholate containing 1.0 mg ml $^{-1}$ PC and centrifuged at 110,000 g for 30 min. After 2–5 h incubation of the supernatant at 4 $^{\circ}$ C, stable high M_r complexes ($>1 \times 10^6$ K) of the cG-mediated channel spontaneously formed ($t_{1/2}$ 4.5 h). These were isolated from other SRDM proteins by fractionation on a Sephadex 4BCL column equilibrated in buffer A containing 1.0% sodium cholate and 1.5–2.0 mg ml $^{-1}$ PC. Activity partitioned with the high M_r channel complex in the void volume. Further purification was achieved by pelleting the high M_r complex and discarding the supernatant. Remaining column fractions containing rhodopsin and reduced amounts of 39K protein, but enriched in other SRDM membrane proteins, were pooled to obtain fractions such as shown in lane c or assayed separately (data not shown). In some experiments, the top 2–3 inches of the column contained concanavalin A (Con A) bound to Sephadex 4B-CL. Void volume fractions contained the 39K protein complex as for Sephadex 4B-CL columns. Remaining fractions which eluted in the absence of α -methylmannoside contained trace amounts of, or no, 39K protein (lane d) but no rhodopsin. Activity was below detectable levels in these latter fractions. Rhodopsin eluted with 0.25 M α -methylmannoside was inactive. Void volume Con-A-Sephadex column fractions reconstituted in the presence of 0.25 M α -methylmannoside contained channel activity. Detergent exposure times and reconstitution conditions were similar for all protein fractions in each experiment. Protein concentrations for reconstitutions were 0.50–1.25 mg ml $^{-1}$, PC was 3.0–4.0 mg ml $^{-1}$ and detergent concentration, 2.0–2.5%. Samples were dialysed for 24–48 h in 10 mM MOPS, 50 mM KCl, 1.0 mM MgCl $_2$, 0.25 mM DTT, pH 7.0. Membranes were centrifuged at 45,000g directly after dialysis or after freezing at -20° C for 10–15 min and thawing. Reconstituted vesicles were loaded with Az III for Ca $^{2+}$ flux assays as in Fig. 1. Yield of 39K protein was 15–30% of that of original SRDM. SDS-polyacrylamide gels were run as in Fig. 4. Coomassie blue stain was used except for lane e (silver stain¹⁴).

dent Ca $^{2+}$ fluxes when reconstituted by detergent dialysis into PC vesicles (see Fig. 3 legend).

The $K_{1/2}$ for cGMP, Hill coefficient, nucleotide specificity, ion-gradient dependence and sensitivity to 1-*cis*-diltiazem in these reconstituted preparations were similar to those described above for SRDM. Flux per mg total protein was 3–8-fold higher than in SRDM starting material. Unlike SRDM, reconstituted

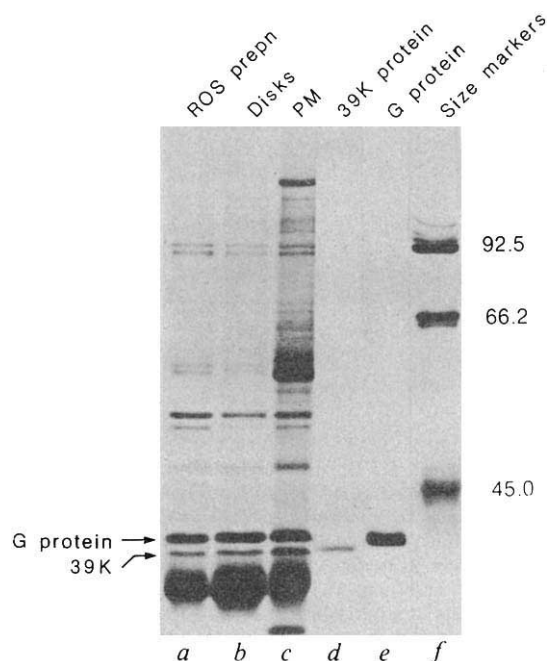


Fig. 4 Proteins from total rod outer segments (ROS), disk-enriched and plasma membrane (PM)-enriched fractions (lanes *a*–*c*), purified 39K protein (lane *d*), purified G protein (lane *e*). **Methods.** ROS with intact plasma membranes were prepared according to Schnetkamp¹⁸ and incubated in buffer A in the presence of Con A-Sepharose beads (Pharmacia). Bound ROS were then hypotonically ruptured to release disks. A column formed from the sedimented beads was washed with 0.25 M α -methylmannoside in 1% octylglucoside which solubilized the membranes and removed Con-A-bound and plasma membrane-associated proteins. SDS-polyacrylamide slab gels were composed of 14.0% acrylamide, 0.12% bis-acrylamide, 0.38 M Tris base, pH 9.0, 0.1% SDS, 25–40 mM KCl, 1.67 μ l ml⁻¹ of 10% ammonium persulphate and 0.4 μ l ml⁻¹ TEMED. Stacking gels were composed of 5.0% acrylamide, 0.1% SDS, 0.1% bis-acrylamide, 0.125 M Tris base, pH 9.0, 25–40 mM KCl, 4.0 μ l ml⁻¹ ammonium persulphate and 1.0 μ l ml⁻¹ TEMED. Gels were run at 10.0 mA per gel for 14–16 h and fixed and stained with Coomassie blue as described by Laemmli²⁹. The 39K protein was purified by soaking gels in 4.0 M sodium acetate for 5–15 min (ref. 30) following electrophoresis. Transparent bands were cut out and proteins eluted from gel slices into 1–3 ml of 0.1% SDS for 24–48 h. G protein was purified by the method of Baehr³¹. The G protein α - and β -subunits usually comigrate in this SDS-polyacrylamide gel system (lane *e*).

vesicles containing the 39K protein/mediated an increased cation flux when 8-bromo-cGMP was trapped in the vesicles. Thus the channel inserts bidirectionally into the membrane during reconstitution. Specific activity of reconstitutions exceeded that of SRDM by up to 16-fold when corrected for 50% inaccessible sites.

Activity was correlated with the amount of 39K protein in various column fractions tested (Fig. 3). Column fractions which did not contain a 39K component were inactive in all cases. Column fractions containing 39K protein and enriched in other SRDM proteins (Fig. 3, lane *c*) were regularly monitored for correlations of activity with the amount of higher M_r proteins. No correlation was found for any of the other SRDM components, because these proteins are enriched (per mg total protein) in lanes *c* and *d*, but activity of these fractions is lower (lane *c*) than control or undetectable (lane *d*). The relative amount of rhodopsin was also not correlated with activity, because its concentration was high in SRDM (>90% of total protein) and lower in reconstituted 39K preparations (~50%), whereas activity per mg total protein increased. Also, purified rhodopsin alone¹⁵ was inactive. Because the 39K protein could

not be freed from rhodopsin under non-denaturing conditions, however, we cannot exclude the possibility that it stabilizes activity. Addition of protease inhibitors (phenylmethylsulphonylfluoride (PMSF), leupeptin, aprotinin, benzamide) during solubilization or reconstitution did not alter the specific activity or the M_r profile of 39K reconstituted preparations. Occasionally, dimers of the 39K preparation could be identified on SDS-polyacrylamide gels. These were stable to reducing agents like dithiothreitol (DTT), could be shown to form from the monomer 39K protein/rhodopsin complex, and had an apparent M_r of 64K. Activity was not correlated with the amount of dimer present on SDS gels. A 250–260K protein was frequently detectable as a minor component (<5%) in our preparations, but we were unable to detect correlation between activity and the amount present on SDS gels. Long detergent exposure times (>10 h) and low KCl content of gel buffers resulted in broadening and poorer resolution of the 39K band from rhodopsin on gels.

We were able to purify the 39K protein further to >95% homogeneity (Fig. 4, lane *d*) for use in amino-acid composition and amino-terminal sequence analyses. Amino-acid analyses indicate that it is neither a proteolytic fragment of G protein, post-translationally modified (for example by phosphorylation) rhodopsin, nor one of the three isoelectric forms of rhodopsin which were reported to have identical amino-acid compositions¹⁶. Amino-terminal sequence analysis indicated that the amino-terminus is not blocked, further distinguishing it from rhodopsin. The first 10 amino acids had no homology with those reported for the three human cone pigments¹⁷.

Gel densitometry shows the 39K protein to be present as $2.5 \pm 1.1\%$ of the rhodopsin content in RDM preparations. This corresponds to a density of approximately $600 \mu\text{m}^{-2}$ for bovine disk membranes. A 39K protein was also detected in *B. marinus* RDM in approximately the same 39K protein/rhodopsin ratio.

Plasma-membrane-dominated preparations were obtained by binding rod outer segments, prepared according to the method of Schnetkamp *et al.*¹⁸, to Con A-Sepharose beads, followed by hypotonic rupture to release the disks. Figure 4 shows analysis of the resultant disk and plasma membrane fractions. Although no definitive marker protein distinguishes plasma from disk membrane, the two membrane fractions are clearly different in composition. Note the enrichment of ~60K and ~110K proteins in the plasma membrane fraction. Ratios of 39K protein to rhodopsin concentrations were higher in plasma membrane fractions by 20–50%, as determined by gel densitometry. However, until it is known whether plasma membranes contain the same concentration of rhodopsin as disk membranes, we cannot determine whether the larger 39K protein/rhodopsin ratio implies enrichment of plasma membrane channel density or reduction of rhodopsin density.

We provide evidence that a 39K protein from bovine RDM mediates cGMP-dependent cation conductance. Our estimated density of $600 \mu\text{m}^{-2}$ is similar to channel densities (100 – $1,000 \mu\text{m}^{-2}$) inferred from noise measurements of amphibian photoreceptor plasma membrane patches^{9,19,20}. The external cGMP binding sites that induce cation flux on our disk preparations correlate topologically with the plasma membrane cytosolic cGMP binding sites studied by excised patch electrophysiology¹. Binding of cGMP, rather than nucleotide hydrolysis, activates both disk and plasma membrane conductances¹ and 1-*cis*-diltiazem inhibits both^{11,21}. We find cooperativity for cGMP with a Hill coefficient (1.7) and $K_{1/2}$ similar to those obtained by Fesenko *et al.*¹ from plasma membrane patch current measurements made at physiological divalent cation concentrations. More recent estimates^{9,20} of the cooperativity index for cGMP in plasma membrane patch current measurements yield the higher value of 3.0, with the $K_{1/2}$ of 9.6–45 μM . These studies were performed at nonphysiologically low divalent cation concentrations (<1 μM) and it has not been determined whether this affects the Hill coefficient and/or binding constant. It is not

yet known whether the 1.7 cGMP cooperativity index for channel activation implies binding of more than one mole of cGMP per mole of 39K protein instead of a 1:1 cGMP/39K protein binding stoichiometry with more than one subunit making up the cation channel. The M_r of other cation channels is generally larger than 39K and multiple identical or distinct subunits are required for activity^{22,23}.

We thank Jeff L. Feldman for help with photography, Ruth Hogue Angeletti and Caryn Glusman for amino-acid and amino-terminal sequence analyses and Boyd Haley for 8-azido-cGMP. This research was supported by grant from the US PHS.

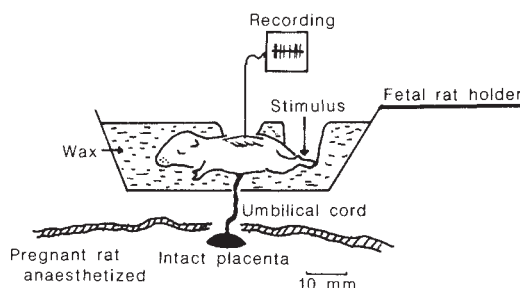


Fig. 1 A diagram of the *in vivo* fetal rat preparation. Hatched double line, surface of mother's abdomen.

Received 29 December 1986; accepted 10 March 1987.

1. Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Nature* **313**, 310–313 (1985).
2. Yau, K.-W. & Nakatani, K. *Nature* **309**, 352–354 (1984).
3. Hodgkin, A. L., McNaughton, P. A. & Nunn, B. J. *J. Physiol., Lond.* **358**, 447–468 (1985).
4. Cohen, A. I., Hall, I. A. & Ferrendelli, J. A. *J. gen. Physiol.* **71**, 595–612 (1978).
5. Lolley, R. N. & Racz, E. *Vision Res.* **22**, 1481–1486 (1982).
6. Polans, A. S., Kawamura, S. & Bownds, M. D. *J. gen. Physiol.* **77**, 41–48 (1981).
7. Torre, V., Matthews, H. R. & Lamb, T. R. *Proc. natn. Acad. Sci.* **83**, 7109–7113 (1986).
8. Zimmerman, A. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8813–8817 (1985).
9. Haynes, L. W., Kay, A. R. & Yau, K.-W. *Nature* **321**, 66–70 (1986).
10. Sitarumayya, A. & Liebman, P. A. *J. biol. Chem.* **258**, 1205–1209 (1983).
11. Stern, J. H., Kaupp, U. B. & McLeish, P. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1163–1167 (1986).
12. Kendrick, N. C. *Analyt. Biochem.* **83**, 433–450 (1977).
13. Steiner, R. A. *et al. Anal. Chem.* **51**, 351–353 (1979).
14. Merrill, C. R. *et al. Science* **211**, 1437–1438 (1981).
15. DeGrip, W. J. *Meth. Enzym.* **81**, 197–207 (1982).
16. Planter, J. J. & Kean, E. L. *Biochem. biophys. Acta* **744**, 312–319 (1983).
17. Nathans, J., Thomas, D. & Hogness, D. *Science* **232**, 193–210 (1986).
18. Schnetkamp, P. P. M., Klompmaekers, A. A. & Daemen, F. J. M. *Biochem. biophys. Acta* **552**, 3379–3389 (1979).
19. Bodia, R. D. & Detwiler, P. B. *J. biol. Chem.* **260**, 6788–6800 (1985).
20. Zimmerman, A. L. & Baylor, D. A. *Nature* **321**, 70–72 (1986).
21. Koch, K. & Kaupp, U. B. *J. biol. Chem.* **260**, 6788–6800 (1985).
22. Barchi, R. L. *Int. Rev. Neurobiol.* **23**, 69–101 (1982).
23. Anhalt, R., Lindstrom, J. & Montal, M. *The Enzymes of Biological Membranes* Vol. 3, 336–401 (Plenum, New York, 1985).
24. Caretta, A. & Cavagioni, A. *Eur. J. Biochem.* **132**, 1–8 (1983).
25. Caretta, A., Cavagioni, A. & Sorbi, R. T. *Eur. J. Biochem.* **153**, 49–53 (1985).
26. Simon, D. *et al. Ann. N.Y. Acad. Sci.* **307**, 52–69 (1978).
27. Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
28. Sedmark, J. J. & Grossberg, S. E. *Analyt. Biochem.* **79**, 544–552 (1977).
29. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
30. Higgins, R. C. & Dahmus, M. E. *Analyt. Biochem.* **93**, 257–260 (1979).
31. Baehr, W., Devlin, M. J. & Applebury, M. L. *J. biol. Chem.* **254**, 11669–11677 (1979).

Spontaneous and evoked activity of fetal primary afferents *in vivo*

Maria Fitzgerald

Department of Anatomy and Embryology, University College London, Gower Street, London WC1 6BT, UK

The first movements of the fetus are apparently random and spontaneous^{1,2}. Their onset coincides with the growth of dorsal root afferents into the spinal cord^{3,4} and it is possible that they are not simply a result of spontaneous motoneuron activity but are reflex responses to sensory stimulation⁵. It is not clear what stimuli could normally evoke such reflexes because nothing is known of the properties of primary afferent neurons in the fetus. I have investigated this by making recordings from single dorsal root ganglion cells in fetal rats *in vivo*. The afferents have small, defined receptive fields and respond to mechanical stimulation of skin or muscle at intensities that might occur *in utero*. Many of them are also chemosensitive. Unlike postnatal afferents they display background activity which peaks at the same age as fetal movements. Repeated stimulation causes long-lasting increases of both background and evoked activity. Such sensory input is likely to have a considerable influence on fetal movements and on the development of spinal cord connections.

Primary afferent activity was recorded from rat fetuses *in vivo*. Pregnant rats were anaesthetized with an intraperitoneal (i.p.) dose of urethane (1.25 g per kg body weight) and one fetus was removed, keeping the umbilical cord intact to maintain contact

with the maternal circulation. The fetus was placed just above the mother's abdomen in a dish which pushed together to leave a slot in the bottom for the umbilical cord (Fig. 1). It was then given an i.p. dose of 0.5 g per kg urethane, paralysed and covered with low melting point wax. Two small troughs were made in the wax to allow access to the lumbar cord and to one hindlimb. The L4 and L5 dorsal root ganglia (DRG) were exposed by laminectomy and covered with warm paraffin oil. Extracellular recordings of single unit activity in the DRG were made with glass-coated tungsten microelectrodes. The skin and deeper tissues of the hindlimb were stimulated by brush, touch or firm pressure. In some cases, noxious thermal (water at 48 °C) and chemical stimuli (mustard oil) were applied. The skin was stimulated electrically at 1–5 mA, 100–500 μ A through pin electrodes. Recordings were made over a period of about two hours following surgery.

Successful recordings were obtained from fetuses on the 16–20th day of gestation (age E16 to E20). At E16, a small number of units could be isolated ($n = 10$, three rats) but only by their spontaneous activity (0.1–2.0 Hz), as they did not respond to natural stimulation. Electrical stimulation of the skin at the ankle did evoke a response in four of these units. From E17 onwards units were easier to isolate and had clear peripheral receptive fields. Cutaneous afferents could be divided into rapidly adapting pressure (RAP), slowly adapting pressure (SAP) and rapidly adapting, low-threshold mechanoreceptor (RAM) units. There was also a clear group of muscle afferents (M) which showed regular maintained firing to muscle stretch and joint movement and whose responses were unaffected by removal of the hindlimb skin. Units were also found which had no apparent receptive field at all (NRF). No RAM units were found until E17. Conduction velocities fell in the range 0.3–0.85 m s^{-1} and there was no correlation between conduction velocity and receptor type. Frequency of firing and total number of impulses per stimulus was low compared to the adult or the newborn^{6,7} but increased with fetal age. Table 1 shows the numbers of different afferent types recorded at each age.

RAP units were the most common fetal afferent, responding with a brief (200–500 ms) burst of impulses to pressure (>0.3 g) on the skin (Fig. 2a–c). They were sensitive to increases in stimulus intensity and pinching the skin could evoke a response

Table 1 Primarily afferent units tabulated according to the receptor type and fetal age

Fetal age	Type of receptor					Total
	RAP	SAP	RAM	M	NRF	
E17	7	2	0	1	2	12
E18	8	3	2	3	1	17
E19	6	4	5	3	1	19
E20	6	3	4	2	1	16

RAP, rapidly adapting pressure receptor; SAP, slowly adapting pressure receptor; RAM, rapidly adapting mechanoreceptor; M, muscle afferent; NRF, no receptor field.

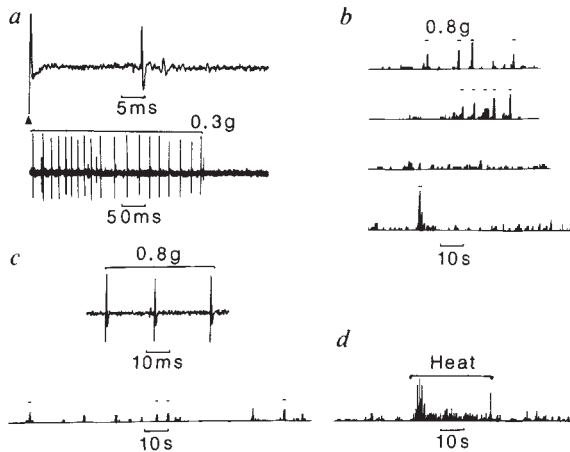


Fig. 2 Examples of fetal RAP receptors. *a*, RAP unit recorded on E18, with a conduction velocity 0.75 m s^{-1} . Top trace, response to electrical stimulation; lower trace, the response to 0.3 g von Frey hair on the receptive field, a spot on the skin of one toe. This was typically a burst of 7–20 impulses lasting 200–400 ms. *b*, Another RAP recorded on E18, with conduction velocity 0.45 m s^{-1} and a threshold of 0.6 g on a small zone (1.0 mm^2) of the upper leg. The four traces are continuous and show the initial background activity in this unit which steadily increases with repeated probing with a 0.8 g von Frey hair (small bars above the trace). In the bottom trace it can be seen that the evoked response is also increased. *c*, RAP unit, recorded on E17, with a conduction velocity of 0.62 m s^{-1} and a threshold of 0.8 g on a zone on the ankle. The unit fired 3–6 impulses over 50–150 ms (upper trace). Response magnitude was typically less than on E18. This afferent had bursting background activity (lower trace) that was very similar to the response evoked by von Frey hair stimulation. *d*, Response to bathing the hindlimb in water at 48°C of an RAP unit recorded on E19. Note the background activity and its increase following the heat stimulus.

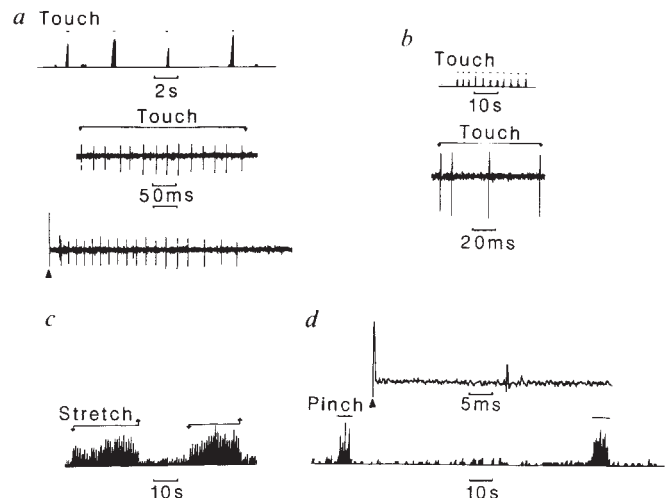


Fig. 3 *a, b*, Examples of fetal touch RAM receptors; *c, d*, receptors of the M and SAP type. *a*, Typical RAM2 receptor recorded on E19, conduction velocity 0.5 m s^{-1} . Top two traces show the very regular burst of impulses at 40 Hz in response to a touch stimulus (0.03 g von Frey hair) on a spot on the sole of the foot (small bars above trace). Bottom trace, the same burst produced by very low intensity ($15 \mu\text{A}$, $50 \mu\text{s}$) electrical stimulation of the skin. *b*, Typical RAM1 receptor recorded on E19. Responses to a 0.03 g von Frey hair were brief, highly repeatable bursts of 3–4 impulses. *c*, Response of a muscle afferent on E19 to flexion of the ankle joint and stretch of the gastrocnemius muscles. *d*, A SAP unit recorded on E18, conduction velocity 0.38 m s^{-1} . Top trace, response to electrical stimulation of the skin; bottom trace, slowly adapting response to pinching the receptive field, a zone of 1 mm^2 on the lower lateral leg. Note the background activity in this afferent.

lasting for up to 2 s. SAP units had higher thresholds ($>0.8 \text{ g}$) and produced a more maintained response of 5–10 s to pressure or pinch of the skin (Fig. 3*d*). Afferents in both the RAP and SAP category were found that responded to heating (Fig. 2*d*) or applying the irritant chemical mustard oil to the skin. RAM afferents fell into two groups. The first (RAM1) had very low-threshold ($<0.08 \text{ g}$), rapidly adapting, highly repeatable responses (Fig. 3*b*). These bore a resemblance to rapidly adapting hair and touch afferents in the adult and newborn^{6,7}. The other group (RAM2, 30% of all RAM cells) gave a very distinctive and highly regular, 20–50 Hz burst of impulses to touching the skin that lasted for up to 0.5 s. The same burst was evoked by low-intensity electrical stimulation of the skin (Fig. 3*a*). Muscle afferents were distinguishable by their ability from E17 to fire for 15–20 s at 10–20 Hz in response to a small change of limb position (Fig. 3*c*).

Two important features of fetal afferents were observed that are not prominent in neonates or adults^{6,7}. Firstly, there was considerable spontaneous activity in fetal DRGs. This could be observed from the first electrode penetration before the limb had been stimulated. It peaked at E18 and E19 when 58% and 37% respectively of cells had background activity with a mean frequency, in the absence of recent skin stimulation, of 0.5–1.5 Hz. On E17 it was observed in 33% of units but was at a very low frequency of $<0.5 \text{ Hz}$. At E20, 21% of units were spontaneously active with a mean frequency of 0.5–2.0 Hz. The activity was irregular and in some cases occurred in high-frequency bursts, with seconds of silence or activity at very low frequency between them. Spontaneous activity was observed in all afferent types except the RAM1 group. Section of the dorsal roots had no effect on the activity. Examples can be seen in Fig. 2*b, c* and *d* and Fig. 3*c* and *d*.

Another feature of fetal afferents was their sensitization on repeated stimulation. Unlike the sensitization described for adult nociceptors⁷, the effect in fetal units was produced with repeated low-intensity mechanical stimulation. It was characterized by a marked increase in background activity and increased evoked responses and was observed in 50% of cells on E18 and E19. Heating the skin was even more effective and could be followed by background activity of 10–20 Hz. Sensitization was observed in all afferent types except the RAM1 group.

These results show that sensory afferents are very active between E16 and E20. Their peripheral terminals clearly have defined receptive fields and respond to cutaneous touch, pressure, heat and chemical stimulation and to limb movement despite the fact that specific sensory end-organs in skin or muscle have not fully differentiated^{8–10}. An important feature of their activity is that it can be related to fetal movements. Firstly the level of background firing in sensory afferents peaks at the same fetal age as spontaneous motility^{1,2}. Secondly, repeated peripheral stimulation increases both their evoked and background firing. This is consistent with the pattern of fetal movements which tend to build up into peaks with periods of inactivity in between. The normal somatic stimuli for sensory afferents *in vivo* may come from uterine contractions or from self-stimulation but a further important point about these results is that they show that there is intrinsic afferent input in the absence of obvious somatic stimulation. This background activity may be partly a result of the fetal environment in these experiments but its pattern and dependence on age argue against this. One factor could be unstable or immature membrane properties of developing ganglion cells¹¹ or alternatively, as fetal afferents are chemosensitive, the activity may be evoked by chemicals in the developing target tissue.

This work was supported by the MRC. I thank Alan Ainsworth and Mary Gardner for technical help and Clifford Woolf & Pat Wall for their advice and encouragement.

Received 20 November 1986; accepted 22 January 1987.

1. Angulo y Gonzalez, A. W. *J. comp. Neurol.* **55**, 395-442 (1932).
2. Narayanan, C. H., Fox, M. W. & Hamburger, V. *Behaviour* **40**, 100-134 (1971).
3. Altman, J. & Bayer, S. A. *The Development of the Rat Spinal Cord* (Springer, Berlin, 1984).
4. Smith, C. L. *J. comp. Neurol.* **220**, 29-43 (1983).
5. Windle, W. F. *The Physiology of the Foetus* (Phildalephia, 1940).
6. Fitzgerald, M. J. *Physiol. Lond.* **383**, 79-92 (1987).
7. Lynn, B. & Carpenter, S. E. *Brain Res.* **238**, 29-43 (1982).
8. Diamond, J. & Nurse, C. A. *Neuroscience* **11**, 509-520 (1984).
9. Zelena, J. *J. Embryol. exp. Morph.* **5**, 283-289 (1957).
10. Bradley, R. M. & Mistretta, C. M. *Physiol. Rev.* **55**, 352-382 (1975).
11. Fulton, B. P. *Neurosci. Lett.* **73**, 125-130 (1987).

Relationship between the nerve action potential and transmitter release from sympathetic postganglionic nerve terminals

James A. Brock & Thomas C. Cunnane

University Department of Pharmacology, South Parks Road, Oxford OX1 3QT, UK

At the skeletal neuromuscular junction, electrophysiological methods have provided much useful information about the mechanisms involved in the release of transmitter¹⁻³. At the autonomic neuroeffector junction it has not been possible to carry out similar studies⁴. Here we report a method of extracellular recording which allows simultaneous measurement of both the nerve action potential and transmitter release from postganglionic sympathetic nerve terminals. We have confirmed that release is intermittent^{5,6}, but the importance of this new approach is that the relationship between the nerve terminal action potential and transmitter release can be studied unambiguously for the first time. Thus we are able to show unequivocally that intermittence is caused by a low probability of release in the invaded varicosity and not by failure of the action potential to invade the varicosity.

Vasa deferentia were removed from male guinea pigs (Duncan Hartley, 350-600 g) and individual preparations pinned to the Sylgard (Dow Corning)-covered base of a 3 ml glass organ bath, which was mounted on the stage of a Zeiss ACM microscope. The organ bath was perfused continuously at 3 ml min⁻¹ with Krebs solution of the following composition (mM): NaCl, 118.4; NaHCO₃, 25.0; NaH₂PO₄, 1.13; CaCl₂, 1.8; KCl, 4.7; MgCl₂, 1.3 and glucose 11.0. The solution was gassed with a mixture of 95% O₂ and 5% CO₂ (pH 7.4) and maintained at 36-37 °C. A window was cut in the connective tissue sheath enveloping the vas deferens, and a bevelled glass electrode (tip diameter <50 µm) filled with Krebs solution of the same composition as the bath solution applied to the muscle surface with slight suction (seal resistance <1 MΩ). Electrical activity was recorded through an a.c. amplifier (Neurolog NL 104, low frequency cutoff 0.1 Hz) and the output filtered between either d.c. and 0.3 kHz or d.c. and 5 kHz (Neurolog NL 125). When action potentials were recorded the high frequency cutoff was set at 5 kHz to avoid distortion of the spike. The vas deferens was electrically stimulated indirectly through the hypogastric nerve trunk. Spontaneous excitatory junction potentials (s.e.j.ps) and excitatory junction potentials (e.j.ps) were also recorded intracellularly using glass micro-electrodes filled with 5 M potassium acetate (resistances 30-60 MΩ)⁶.

In the absence of stimulation, spontaneous negative-going potentials were recorded in 130 attachments in 34 preparations. Results from 5 representative attachments are summarized here. First, the amplitude of individual events varied between 500 µV and the noise level of the recording system (maximum 10 µV peak-to-peak) and had an amplitude distribution skewed towards the low-amplitude events. The mean amplitude was 22 ± 11 µV (mean ± s.d.) *n* = 745 events. Second, the mean

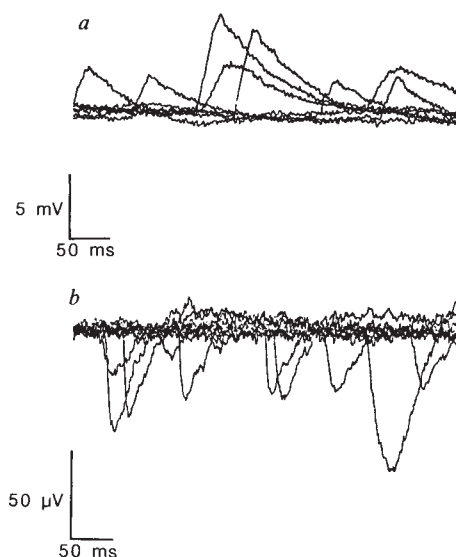


Fig. 1 *a*, Spontaneous excitatory junction potentials (s.e.j.ps) recorded intracellularly from the guinea pig isolated vas deferens. *b*, Spontaneous excitatory junction currents (s.e.j.cs) recorded extracellularly. The s.e.j.ps and s.e.j.cs were digitized at 2 kHz and a series of random traces superimposed photographically to show the characteristic features of these events.

frequency of occurrence was 0.21 ± 0.04 Hz. Third, the average 10-90% rise time was 8.3 ± 5.6 ms (*n* = 745) and the time to half decay was 17.4 ± 7.4 ms (*n* = 745) with a total duration of 50-80 ms. Fourth, they were unaffected by α -adrenoceptor antagonists but were blocked by α, β -methylene adenosine triphosphate which desensitizes P₂ purinoceptors⁷. Fifth, these signals were tetrodotoxin resistant suggesting that they were not caused by spontaneously occurring, sodium-dependent nerve action potentials, and were abolished when the sympathetic nerves were destroyed by pretreatment of the animals with 6-hydroxydopamine (day 1, 150 mg per kg; day 2, 250 mg per kg; day 3, vasa removed). These characteristics suggest that they are the extracellular equivalent of the intracellularly recorded s.e.j.p.^{8,9} and therefore reflect the release of single quanta from varicosities under the suction electrode. By analogy with focal extracellular recording at the skeletal neuromuscular junction^{1,2}, there is good reason to believe that these potentials are a measure of the time course of the current underlying the s.e.j.p. For this reason, they have been termed spontaneous excitatory junction currents (s.e.j.cs). Figure 1*b* shows a series of s.e.j.cs recorded from a single attachment to the guinea pig vas deferens. It is clear from separate studies that the mean total duration of the s.e.j.p. (time constant of decay 53 ms) is approximately twice that of the s.e.j.c. (time constant of decay 25 ms) (Fig. 1).

Three patterns of postjunctional electrical activity were observed in response to nerve stimulation. The first pattern was recorded when the stimulus intensity was slowly raised during continuous stimulation at 1 Hz, and consisted of negative-going potentials which were stimulus locked, and had both an amplitude and a time course similar, and sometimes identical, to those of the s.e.j.cs in the same attachment. At all frequencies of stimulation (0.1-4 Hz) the evoked events at a constant latency occurred intermittently. These brief negative-going events have been termed excitatory junction currents (e.j.cs) (see Fig. 2*b*). The characteristics of these potentials are similar to those of the intracellularly recorded discrete events previously described in the same tissue^{5,6} and we believe that they represent the intermittent, monoquantal release of transmitter from individual varicosities located within the recording electrode. These characteristics of transmitter release from peripheral autonomic nerves thus closely resemble those in central synaptic boutons^{10,11}. The

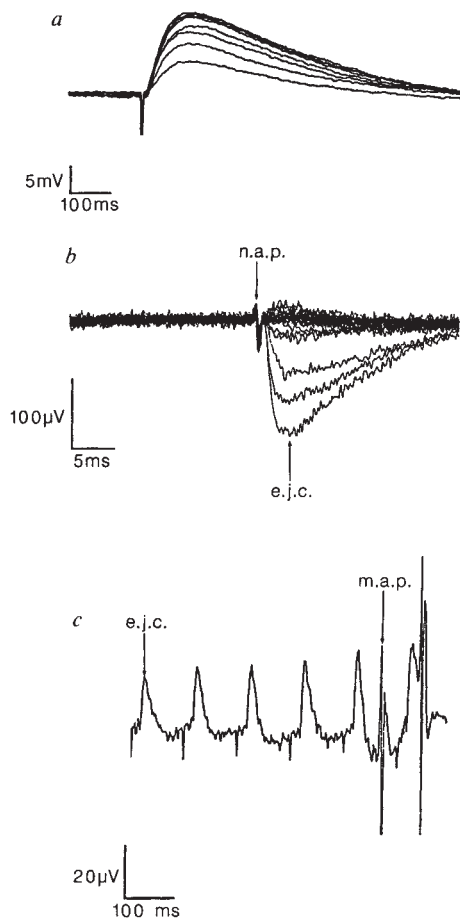


Fig. 2 Patterns of electrical activity recorded intracellularly and extracellularly from the guinea pig vas deferens following stimulation of the hypogastric nerve. *a*, The e.j.ps were recorded with an intracellular microelectrode (7 pulses at 1 Hz) showing the characteristic slow time course of the e.j.p. and facilitation. *b*, Simultaneous measurement of the nerve terminal action potential (n.a.p.) and evoked e.j.cs recorded with a suction electrode (continuous train of 25 pulses at 1 Hz) showing that release is intermittent. *c*, The e.j.cs exhibiting facilitation, but no summation (train of 5 pulses at 5 Hz), and triggering an action potential in the smooth muscle (m.a.p.). The negative-going deflection preceding each e.j.c. is the stimulus artefact.

second pattern of electrical activity was usually observed when the stimulation intensity was further increased, and differed from the first pattern in that there was a non-intermittent, superimposed positive-going potential, which displayed frequency-dependent facilitation and was graded with stimulus strength. We interpret these positive-going potentials to be the extracellular equivalent of the e.j.p., involving transmitter release from many varicosities, the sign of the potential indicating that current was flowing into the suction electrode as a result of transmitter-induced depolarization of cells adjacent to the site of attachment. The third pattern of postjunctional activity recorded from the surface of the vas deferens was observed when the frequency of stimulation was increased from 1 Hz to 2–5 Hz. Under these conditions, the e.j.cs could trigger a negative or positive brief potential associated with contraction of the vas which we have interpreted as the smooth muscle action potential (Fig. 2*c*). Note that no summation of the e.j.cs was observed at frequencies of stimulation up to 5 Hz; this finding further emphasizes the brief nature of these events compared to e.j.ps (Fig. 2*a*) which normally summate when evoked by stimulation frequencies greater than 1 Hz (ref. 8).

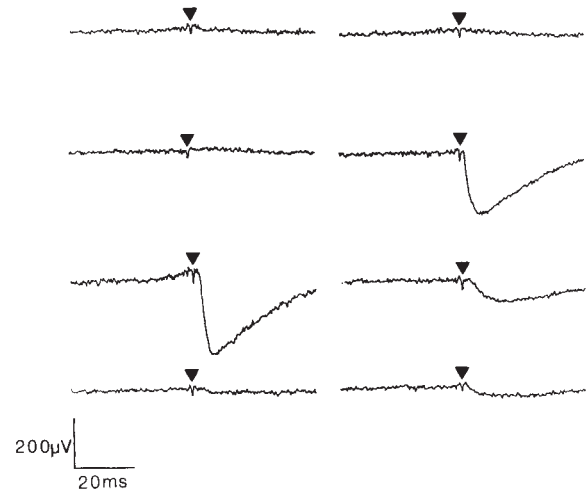
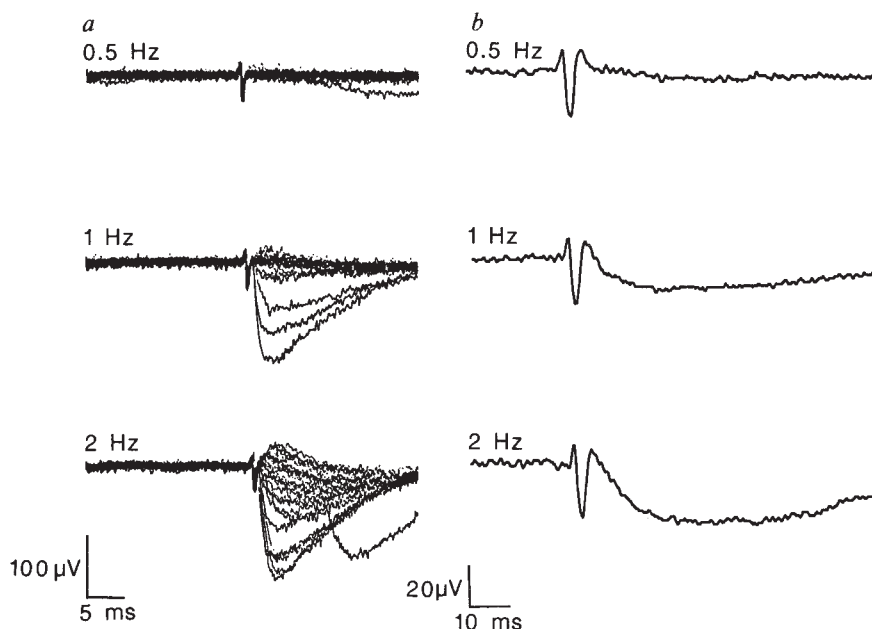


Fig. 3 A series of eight traces showing the relationship of the nerve action potential (n.a.p.) to transmitter release in the guinea pig vas deferens. The n.a.ps (▼) and e.j.cs were evoked by stimulation of the hypogastric nerve with suprathreshold stimuli at 1 Hz. Although the nerve terminal action potential arrived in the secretory terminals on every occasion, the transmitter release process was not always activated. Intermittence of transmitter release was therefore not caused by failure of the nerve terminal action potential to invade the varicosities but resulted from a low probability of release in the invaded varicosity.

We have also studied, on a pulse-by-pulse basis, the relationship between the action potential in, and transmitter release from, the secretory terminals to distinguish between the two possible causes of intermittence of transmitter release, namely either that the nerve action potential invades the varicosity every time the parent axon is stimulated but the probability of transmitter release is low, or alternatively that the nerve action potential fails to reach the varicosity¹². It is now clear that although the nerve action potential always arrives at the varicosity, release is still intermittent (Fig. 3). The action potentials investigated were probably single units because they occurred in an all-or-none manner and during repetitive stimulation they retained the same shape and could not be broken down into separate components. We have also demonstrated by electron microscopy that single varicose fibres run on the surface of the vas deferens. The link between the arrival of the nerve action potential and transmitter release was established by deliberately setting the stimulus strength close to threshold, so that the action potential evoked by trains of several hundred stimuli arrived 'intermittently'; on every occasion that an e.j.c. was elicited it was preceded by a nerve action potential (delay 1–3 ms). Furthermore, variation in the size of the action potential cannot account for failure of release. Apparent variation in the size of the spike (see Fig. 3) can be accounted for by variance in the background noise. It is important to note that there was no detectable difference between averaged action potentials which evoked transmitter release, and those which did not.

It can be clearly demonstrated that transmitter release is dependent on the frequency of stimulation (see Fig. 4*a*). One possible explanation for facilitation of transmitter release from sympathetic nerves would be a frequency-dependent increase in the size or duration of the nerve terminal action potential¹³. Increasing the frequency of stimulation increased the probability of observing evoked e.j.cs but there was no concomitant alteration in the size or shape of the nerve terminal action potential (Fig. 4*b*). Therefore, the increased probability of release associated with facilitation is due to some other mechanism acting at the level of the individual release site. Whether facilitation is due to an enhanced probability of transmitter release from the

Fig. 4 Frequency-dependent facilitation of transmitter release in the guinea-pig vas deferens. *a*, Extracellular recording of nerve action potentials (n.a.ps) and excitatory junction currents (e.j.cs) evoked by stimulation of the hypogastric nerve with trains of 25 suprathreshold stimuli at 0.5, 1 and 2 Hz. *b*, Averages of 25 n.a.ps and associated e.j.cs in another attachment, showing that the facilitation is not associated with a detectable alteration in the configuration of the n.a.p. It is clear that the shape of the n.a.p. is the same whether or not release occurs (for example, compare 0.5 Hz with 2 Hz).



same varicosity or more varicosities releasing transmitter remains to be established.

In conclusion this technique overcomes many of the difficulties associated with intracellular recording from smooth muscle cells and shows directly that (1) the current underlying the e.j.p. is brief compared to the total duration of the e.j.p.; (2) transmitter release from sympathetic nerve terminals occurs intermittently following nerve stimulation; (3) transmitter release is intermittent because of a low probability of release in the invaded varicosity and (4) facilitation is not associated with a detectable change in the configuration of the nerve terminal action potential. This new approach to the study of autonomic transmission will allow the precise site of action of many therapeutic agents to be determined, and also permit a detailed study of the mechanisms involved in the local regulation of transmitter release.

This work was supported by the British Heart Foundation and the MRC.

Received 11 November 1986; accepted 4 February 1987.

1. Fatt, P. & Katz, B. *J. Physiol., Lond.* **117**, 109-128 (1952).
2. Del Castillo, J. & Katz, B. *J. Physiol., Lond.* **132**, 630-649 (1956).
3. Katz, B. *The Release of Neural Transmitter Substances* (Liverpool University Press, 1969).
4. Holman, M. E. in *Smooth Muscle* (eds Bülbüling, E., Brading, A. F., Jones, A. W. & Tomita, T.) 244-288 (Arnold, London, 1970).
5. Blakeley, A. G. H. & Cunnane, T. C. *J. Physiol., Lond.* **296**, 85-96 (1979).
6. Cunnane, T. C. & Stjärne, L. *Neuroscience* **13**, 1-20 (1984).
7. Sneddon, P. & Burnstock, G. *Eur. J. Pharmac.* **100**, 85-90 (1984).
8. Burnstock, G. & Holman, M. E. *Pharmac. Rev.* **18**, 481-493 (1966).
9. Allcorn, R. J., Cunnane, T. C. & Kirkpatrick, K. *Br. J. Pharmac.* (in the press).
10. Jack, J. J. B., Redman, S. J. & Wong, K. *J. Physiol., Lond.* **321**, 65-96 (1981).
11. Horn, A., Triller, A., Mallet, A. & Faber, D. S. *Science* **213**, 898-901 (1981).
12. Alberts, P., Bartfai, T. & Stjärne, L. *J. Physiol., Lond.* **312**, 297-334 (1981).
13. Brown, G. L. & Holmes, O. *Proc. R. Soc. B* **145**, 1-14 (1956).

Fibronectin receptor structures in the VLA family of heterodimers

Yoshikazu Takada, Christina Huang
& Martin E. Hemler

Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA

Multiple cell surface proteins of relative molecular mass 115,000-155,000 (M_r 115K-155K) have been implicated as receptors mediating adhesion to extracellular matrix proteins¹⁻¹⁶. But the organization and relatedness of these peptides has remained unclear. In separate studies, the 'very late antigens' VLA-1 (M_r 210K/130K) and VLA-2 (M_r 160K/130K) were initially characterized as surface heterodimers appearing 2-4 weeks after *in vitro* stimulation of human T cells¹⁷⁻¹⁹. Three more VLA heterodimers have since been discovered, which, like VLA-1 and VLA-2 (refs 18, 20, 21), are each composed of unique α -subunits in association with a common 130K β subunit²². This paper shows that the common VLA β -subunit is equivalent to subunits found in structures with known fibronectin and laminin receptor activity⁹⁻¹², and that VLA-3 and VLA-5²² are similar or identical to these previously defined receptors for adhesion molecules. Antibody blocking studies confirmed that at least some of the widely distributed VLA proteins of previously unknown function are involved in cell

adhesion to fibronectin and laminin. We suggest that the VLA family of receptors may provide cells with multiple independent substrate adhesion capabilities.

Antibody blocking experiments were used to determine if VLA proteins mediate cell adhesion to matrix proteins. Rabbit antiserum to the VLA β -subunit specifically prevented the adhesion of MG-63 cells to plastic coated with fibronectin or laminin (Fig. 1A, C).

To see if VLA proteins resemble previously identified fibronectin and laminin receptor complexes, rabbit antisera were used for immunochemical comparisons. Rabbit anti-chicken fibronectin/laminin receptor (cFNR)¹⁰⁻¹² sera immunoprecipitated multiple proteins (Fig. 2A) from human cell lines, that were similar to VLA proteins (lanes c-f). Rabbit anti-cFNR sera analysed on chick embryo fibroblasts (CEF, lane g) yielded three bands as in previous reports¹⁰⁻¹³, and the lower two bands (known as band 2 and band 3) aligned with those obtained from human cells. The monoclonal anti-human VLA antibody A-1A5 did not crossreact with CEF (lane h). Upon reduction, the α -band from human MG-63 cells seen by cFNR sera comigrated with β (lane a), whereas the 160K VLA α^2 band remained distinct from β (lane b). Rabbit antisera against purified human fibronectin receptor (hFNR)^{9,23} also immunoprecipitated VLA β -like and α -like subunits (Fig. 2B, lanes a, c, e) that resembled the VLA proteins (lanes b, d, f) from each cell line tested. Although multiple α -subunits appear in anti-FNR precipitations (Fig. 2A,B), it was not clear if these subunits were directly recognized, or indirectly coprecipitated with β .

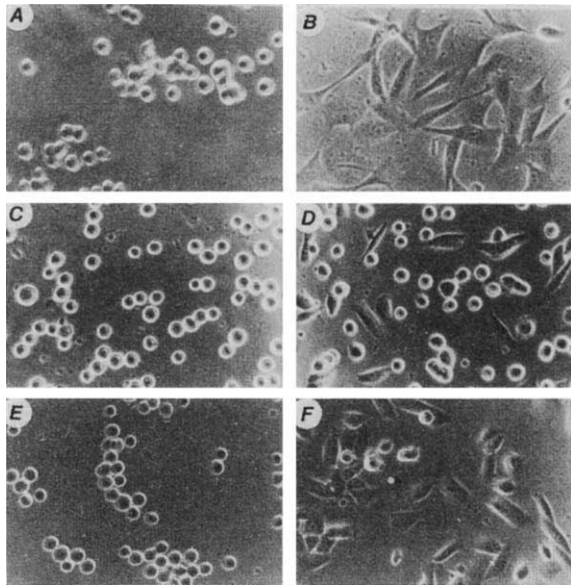


Fig. 1 Blocking of cell adhesion to fibronectin and laminin by anti-VLA antisera. Microtitre wells were pre-coated with fibronectin (A, B), laminin (C, D) or BSA (E), and then MG-63 (human osteosarcoma) cells were plated in the presence of 30-fold diluted anti-VLA rabbit serum (A, C) or control rabbit serum (B, D) or no serum (E). Also, anti-VLA serum was added to cells plated on tissue culture plastic without protein pre-coating (F). Rabbit anti-VLA sera inhibited cell binding to fibronectin (A) and to laminin (C); no serum (not shown) or control serum (B, D) had no effect. There was no adhesion to BSA-coated plates (E) or to uncoated plates (not shown), confirming that cell adhesion required matrix protein. Cell adhesion to tissue-culture plastic (which does not require added matrix protein) was not blocked by anti-VLA β sera (F), supporting the specificity of blocking for matrix-protein-mediated adhesion. The specificity of this anti-VLA β serum was confirmed both by immunoblotting (Fig. 3, lanes l, m, n, o), and by immunoprecipitation (Fig. 4A, lane c). Other preparations of anti-VLA β antisera and other cell lines gave comparable blocking results to those in Fig. 1.

Methods. Microtitre plates (96-well, Linbro Titertek), were incubated with $10 \mu\text{g ml}^{-1}$ fibronectin (Sigma), laminin (Collaborative Research) or bovine serum albumin (BSA) in phosphate-buffered saline (PBS) for 12 h at room temperature. After wells were washed, recently trypsinized MG-63 cells (human osteosarcoma, ATCC) were plated in alpha media with 10% fetal calf serum (FCS), and incubated at 37°C in the presence or absence of rabbit antiserum. Unbound cells were not washed away before cells were photographed (2–3 h after plating). Anti-VLA serum was obtained from a rabbit immunized with native VLA proteins immunopurified from the T-cell line HPB-ALL (as described elsewhere), and reacted primarily with the VLA β -subunit. Weak anti- α^4 reactivity was also present, but inconsequential because MG-63 cells are VLA-4 negative.

To analyse fibronectin receptor crossreactivity further, and to determine which VLA subunits were recognized by anti-FNR sera, we carried out immunoblotting experiments (Fig. 3). Under nonreducing conditions, anti-hFNR sera (lanes b–f), the anti-VLA monoclonal antibody (mAb) TS2/16 (lanes g–k), and anti-VLA β rabbit serum (lanes l–o) all bound to β -subunit material, yielding nearly identical results. Although the cells in this panel express all five VLA α -subunits, no blotting of α -subunits by anti-hFNR sera was observed, suggesting that anti- α antibodies were either not present, or too low in affinity owing to the denaturing conditions.

Another approach to discerning which α -subunits might be directly recognized by anti-cFNR or anti-hFNR sera involved immunoprecipitations of dissociated α - and β -subunits. Both anti-hFNR serum and anti-cFNR serum recognized purified VLA β -subunit (Fig. 4A, lanes d, e; Fig. 4B, lanes b, c). But in

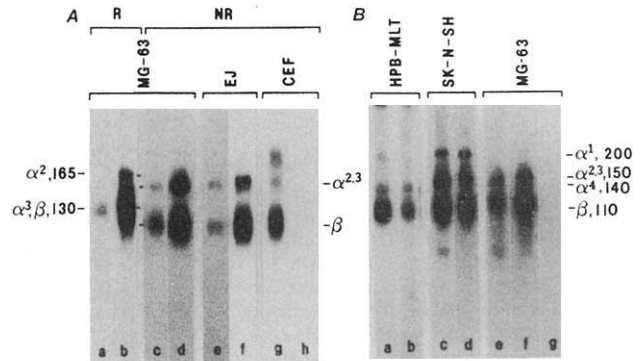


Fig. 2 Comparison of proteins immunoprecipitated by anti-VLA and anti-fibronectin receptor reagents. A, Extracts from the ^{125}I -labelled human osteosarcoma cell line MG-63, the human bladder tumour line EJ and cultured chick embryo fibroblasts (CEF) were immunoprecipitated using the anti-VLA β monoclonal antibody A-1A5^{17,21} (lanes b, d, f, h) and rabbit anti-fibronectin/laminin receptor (cFNR) from chicken (lanes a, c, e, g). Results were analysed using SDS-PAGE (7% acrylamide) with (lanes a, b) or without reduction (lanes c–h). B, Extracts from the ^{125}I -labelled T leukaemic line HPB-MLT, the neuroblastoma line SK-N-SH, and the osteosarcoma line MG-63 were immunoprecipitated using anti-human fibronectin receptor (hFNR) (lanes a, c, e) or A-1A5 (lanes b, d, f) or control sera (lane g) and analysed using nonreducing conditions.

Methods. Radiolabelling with ^{125}I , cell extraction with NP-40, and immunoprecipitation were carried out as described^{21,22}. The anti-cFNR reagent was kindly provided by Dr K. Yamada (NIH) and had been prepared against chick embryo material immunopurified using the mAb J2E as described^{11,12}. The anti-hFNR reagent was kindly provided by Dr M. Pierschbacher (La Jolla) and had been prepared against affinity-purified fibronectin receptor from human placenta as described^{9,23}. The α -subunits expressed by the different human cell lines were identified by antibody recognition, size and two-dimensional gel patterns as described²².

Fig. 4A, anti-hFNR (lane d) but not anti-cFNR serum (lane e) detected the α^5 subunit, whereas in Fig. 4B, anti-cFNR (lane c) but not anti-hFNR (lane b) detected the α^3 -subunit. When all VLA antigens (Fig. 4C, lane a) were removed from a cell extract by preclearing (lane c), all hFNR antigens (lane b) were also removed (lane d), emphasizing the identity or cross-reactivity between VLA proteins and fibronectin receptor structures seen in Fig. 4A,B.

The α^3 and β subunits of human VLA-3 are antigenically similar to part of the avian complex defined by the mAb CSAT and J2E which has three distinct proteins and both laminin and fibronectin binding activities^{10–13}. Although VLA proteins may bind both fibronectin and laminin (see Fig. 1), it is not yet clear if VLA-3 in particular is involved in both activities. As the VLA-3 heterodimer appears to be analogous to only part of the avian complex (band 2/band 3), it may be involved in only one of the functions. Fibronectin and laminin binding have been separated in other studies^{6,9,15}. Perhaps VLA-3 plus another VLA heterodimer (possibly analogous to avian band 1/band 3) accounts for both activities, as VLA proteins are involved in both fibronectin and laminin binding. It is not yet clear which, if any, VLA component might be analogous to avian band 1.

VLA-3 and VLA-5 both have M_r of 150K/110K nonreduced and 135K/130K reduced²², and thus closely resemble the structures previously reported for human and mouse fibronectin receptors^{9,14}. But if VLA-3 is a fibronectin receptor, it is different from the VLA-5-like fibronectin receptor complex isolated from human placenta^{9,23}. The separate recognition of α^3 - and α^5 -subunits by anti-cFNR and anti-hFNR sera respectively provides further evidence that VLA-3 and VLA-5 have structurally distinct α -subunits, consistent with previous results²².

Fig. 3 Comparison of anti-VLA and anti-fibronectin receptor reagents by immunoblotting. Proteins from unlabelled cell extracts were separated by SDS-PAGE using nonreducing conditions and then transferred to nitrocellulose sheets for immunoblotting as described²². After incubation with rabbit anti-hFNR serum (lanes b-f), the monoclonal antibody TS2/16¹⁸ (anti-VLA β , lanes g-k), or rabbit anti-VLA β serum (lanes l-o, used in Fig. 1), nitrocellulose pieces were incubated with ¹²⁵I-labelled goat anti-rabbit or anti-mouse antibody and visualized by autoradiography. Cell lines were chosen which express mostly VLA-1 (SK-N-SH), VLA-2 (GM 3349), VLA-3 (MG 63), VLA-4 (HPB-MLT), or VLA-5 (K-562). As a standard, immunopurified β subunit was directly radiolabelled and then transferred to nitrocellulose with no subsequent incubation with antibody (lane a). Each antibody preparation recognized the 110K VLA β -subunit, as well as the 90K internal form²² of the β -subunit in varying amounts from each of the four or five cell sources (lanes b-o).

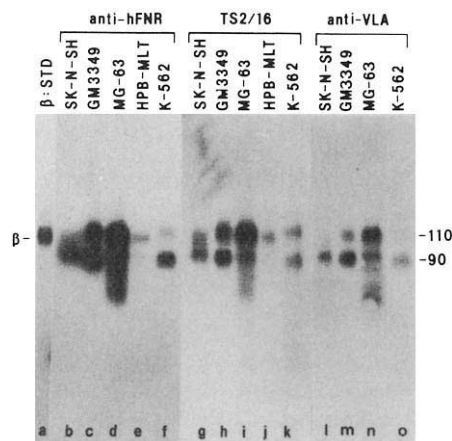
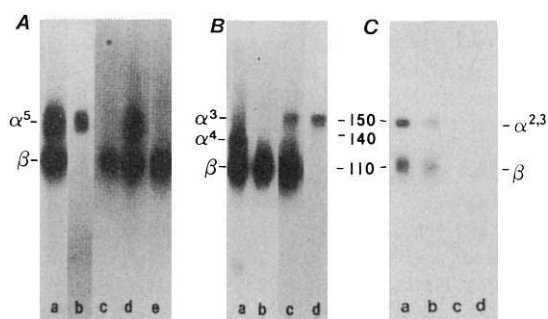


Fig. 4 A, B, Immunoprecipitation of immunopurified, dissociated VLA components using anti-cFNR, anti-hFNR, and anti-VLA reagents. C, Preclearing of hFNR antigens by anti-VLA β reagents. A, Purified VLA-5 was obtained from K-562 cells by A-1A5-Sepharose immunoaffinity chromatography and, after ¹²⁵I-radiolabelling, was analysed without immunoprecipitation (lane a), or immunoprecipitated with mouse anti-VLA-5 (lane b), anti-VLA β (lane c), anti-hFNR (lane d), or anti-cFNR (lane e). A sample of purified VLA-5 contained both α^5 and β (lane a). Evidence for dissociation is that α^5 alone (lane b) and β alone (lane c) could be independently immunoprecipitated by different anti-VLA sera. B, Immunoaffinity-purified, ¹²⁵I-radiolabelled VLA material from MOLT-4 cells was immunoprecipitated using anti-VLA-4 rabbit sera (lane a), anti-hFNR (lane b), anti-cFNR (lane c), or the anti- α^3 mAb J143 (ref. 22; lane d). The VLA proteins purified from the cell line MOLT-4 contained α^4 and β as the major proteins recognized by anti-VLA-4 sera (lane a). Whereas anti-hFNR sera bound only the β -subunit (lane b), anti-cFNR sera precipitated the β -subunit plus a small amount of α^3 subunit (lane c), whereas the mAb J143 (lane d) recognized α^3 alone. Because α^3 and β were dissociated in this preparation, anti-cFNR (lane c) but not anti-hFNR (lane b) is presumed to bind α^3 directly. For A and B, the conditions of high pH (11.5) used for elution of immunoaffinity columns probably cause the observed dissociation of VLA protein subunits. Neither anti-hFNR nor anti-cFNR recognized α^4 (Part B), or recognized dissociated α^1 and α^2 from other VLA preparations (not shown). C, Cell extract was prepared from ¹²⁵I-labelled SK-N-SH cells. After mock-preclearing using negative control antibodies (lanes a, b) or preclearing with a mixture of A-1A5 monoclonal antibody and anti-VLA β serum (lanes c, d), the remaining radiolabelled extract was immunoprecipitated using A-1A5 (lanes a, c) or anti-hFNR serum (lanes b, d).



The expression of VLA-5 on K-562 cell is consistent with reports that K-562 and other erythroleukaemic cells express structures in the M_r region 140K-145K^{14,15} involved in fibronectin binding. The monocytoid cell line U-937 also expresses VLA-5 (ref. 22), a finding that could relate to the discovery of fibronectin receptors on monocytes⁸.

Receptor recognition of matrix proteins such as fibronectin and laminin is directed towards regions of those molecules containing the amino-acid sequence Arg-Gly-Asp (R-G-D)^{9,10,14,15,24-26}. Because anti-VLA β sera blocked binding to fibronectin and laminin, R-G-D recognition may involve a site on the common VLA β subunit, and thus VLA heterodimers may have related functions. In this regard, a function for VLA-1 in cell adhesion may be consistent with the increased VLA-1 expression on T cells present in the normal lung²⁷ or in the rheumatoid synovium²⁸.

Although receptors for certain proteins with R-G-D sequences may be part of the VLA family of heterodimers, it is known that other receptor complexes with R-G-D adhesion function are distinct from VLA proteins. For example, the platelet IIb/IIIa R-G-D recognition molecule and the receptor for the vitronectin R-G-D sequence are different in size from VLA/FNR structures²³, and are not crossreactive with anti-VLA sera (not shown). In summary, placing two different cell adhesion receptors in the context of the VLA family of heterodimers provides a critical step towards understanding the biochemistry of those and related structures. Furthermore, the number of VLA heterodimers (at least five), and their wide distribution, suggest that they have a central role in mediating multiple types of cell adhesion.

We thank Drs K. Yamada and M. Pierschbacher for antisera

against chicken and human adhesion structures. This work was supported by NIH grant CA 42368 and an Arthritis Foundation Investigator Award to M.E.H.

Received 11 December 1986; accepted 29 January 1987.

- Knudsen, K. A., Rao, P. E., Damsky, C. H. & Buck, C. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6071-6075 (1981).
- Tarone, G., Galetto, G., Prat, M. & Comoglio, P. M. *J. Cell Biol.* **94**, 179-186 (1982).
- Neff, N. T. *et al. J. Cell Biol.* **95**, 654-666 (1982).
- Greve, J. M. & Gottlieb, D. I. *J. cell Biochem.* **18**, 221-229 (1982).
- Chapman, A. E. *J. cell Biochem.* **25**, 109-121 (1984).
- Brown, P. J. & Juliano, R. L. *Science* **228**, 1448-1451 (1985).
- Giancotti, F. G., Tarone, G., Knudsen, K., Damsky, C. & Comoglio, P. M. *Exp. Cell Res.* **156**, 182-190 (1985).
- Hosein, B. & Bianco, C. *J. exp. Med.* **162**, 157-170 (1985).
- Pytela, R., Pierschbacher, M. D. & Ruoslahti, E. *Cell* **40**, 191-198 (1985).
- Horwitz, A. F., Duggan, K., Greggs, R., Decker, C. & Buck, C. *J. Cell Biol.* **101**, 2134-2144 (1985).
- Chen, W.-T., Hasegawa, E., Hasegawa, T., Weinstock, C. & Yamada, K. M. *J. Cell Biol.* **100**, 1103-1114 (1985).
- Hasegawa, T., Hasegawa, E., Chen, W.-T. & Yamada, K. M. *J. Cell Biochem.* **28**, 307-318 (1985).
- Knudsen, K. A., Horwitz, A. F. & Buck, C. A. *Exp. Cell Res.* **157**, 218-226 (1985).
- Patel, V. P. & Lodish, H. F. *J. Cell Biol.* **102**, 449-456 (1986).
- Giancotti, F. G., Comoglio, P. M. & Tarone, G. *J. Cell Biol.* **103**, 429-437 (1986).
- Tamkun, J. W. *et al. Cell* **46**, 271-282 (1986).
- Hemler, M. E., Ware, C. F. & Strominger, J. L. *J. Immun.* **131**, 334-340 (1983).
- Hemler, M. E. *et al. J. Immun.* **132**, 3011-3018 (1984).
- Hemler, M. E., Jacobson, J. G., Brenner, M. B. & Strominger, J. L. *Eur. J. Immun.* **15**, 502-508 (1985).
- Pischel, K. D., Hemler, M. E., Huang, C., Bluestein, H. G. & Woods, V. J. *Immun.* **138**, 228-232 (1987).
- Hemler, M. E., Jacobson, J. G. & Strominger, J. L. *J. biol. Chem.* **260**, 15246-15252 (1985).
- Hemler, M. E., Huang, C. & Schwarz, L. *J. biol. Chem.* (in the press).
- Pytela, R., Pierschbacher, M. D., Ginsberg, M. H., Plow, E. F. & Ruoslahti, E. *Science* **231**, 1559-1562 (1986).
- Pierschbacher, M. D. & Ruoslahti, E. *Nature* **309**, 30-33 (1984).
- Akiyama, S. K., Yamada, S. S., Yamada, K. M. *J. Cell Biol.* **102**, 442-448 (1986).
- Wright, S. D. & Meyer, B. C. *J. exp. Med.* **162**, 762-767 (1985).
- Saltini, C., Hemler, M. E. & Crystal, R. *Clin. Immun. Immunopath.* (submitted).
- Hemler, M. E., Glass, D., Coblyn, J. S. & Jacobson, J. G. *J. clin. Invest.* **78**, 696-702 (1986).

Cloning of HTLV-4 and its relation to simian and human immunodeficiency viruses

Hardy Kornfeld, Norbert Riedel, Gregory A. Viglianti, Vanessa Hirsch & James I. Mullins*

Department of Cancer Biology, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115, USA

Although much is now known of the strain variation among the type-1 human immunodeficiency virus (HIV-1), which is the cause of AIDS (acquired immune deficiency syndrome) in the United States, Europe, and Central Africa, much less is yet known about a second group of viruses that have been found in West Africans. One member of this group, named human T-cell lymphotropic virus type 4 (HTLV-4), has been isolated from healthy Senegalese (ref. 1). Another is the virus isolated from West Africans with AIDS-like illness and originally called LAV-2 but now renamed HIV-2 (refs 2 and 3). Both these viruses seem to be less closely related to HIV-1 than they are to a virus of healthy African green monkeys^{4,5}, known variously as simian T-cell lymphotropic virus type 3 (STLV-3) or simian immunodeficiency virus (SIV), which in turn is related to viruses isolated from healthy sooty mangabeys^{6,7} and captive macaques^{8,9-10} with a form of immunodeficiency (to distinguish these viruses they are referred to as STLV-3 (or SIV)_{agm}, STLV-3_{mac} or STLV-3_{smm}). To clarify the relationship between the various HIVs, STLV-3s and HTLV-4 we are determining and comparing the molecular and biological characteristics of several of them. Following our recent publication¹¹ of a restriction-site map of STLV-3_{agm}, we now report that the equivalent map of three isolates of HTLV-4 is remarkably similar to it. In addition we present comparative sequence data on the long terminal repeats (LTR) of HTLV-4, STLV-3_{agm}, HIV-1 and HIV-2, together with evidence that cloned HTLV-4 uses the same receptor as HIV-1 and induces some, but not all, of the cytopathic effects attributed to most isolates of HIV-1 and HIV-2.

Our molecular cloning of STLV-3_{agm} was made possible by a weak nucleic acid hybridization with HIV-1 (ref. 11). This was detected across their entire genomes, but no hybridization was detected between STLV-3_{agm} and HTLV-1 or bovine, feline and sheep retroviruses, indicating that STLV-3_{agm} is one of the closest known relatives of HIV-1.

Using STLV-3_{agm} DNA probes, we have now addressed the genetic relationship between five individual isolates of STLV-3_{agm} (ref. 5), one isolate made following *in vivo* passage of STLV-3_{agm} through a macaque (ref. 5; and P. J. Kanki *et al.*, manuscript in preparation), three isolates of STLV-3_{mac} derived from macaques at the New England Regional Primate Research Center⁹, and three isolates of HTLV-4 (ref. 1). Each virus was grown in HUT-78 cells, a human CD4⁺ T-cell line negative for STLV-3 antigens⁵ and sequences (ref. 11; Fig. 1). In addition, four different proviral DNA clones were obtained from one of the HTLV-4-infected cell lines (PK289).

Our results, summarized in Figs 1 and 2A, show that the restriction site maps of the major form of viral DNA present in each of the six African green monkey cultures and two of the three HTLV-4 isolates were identical at 32 sites. One HTLV-4 isolate (PK289) differed at a single *Pvu*II site. Each of the molecular clones derived from PK-289 cells also displayed this polymorphism. One macaque isolate (number 251) was identical to the consensus whereas the other two (isolates 142 and 239) were identical to each other and differed from the consensus at 2 out of 23 sites mapped (Fig. 2A). The patterns observed were consistent with the absence of one *Pst*I site near the 3' end and

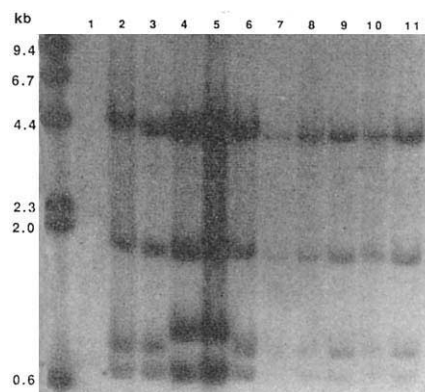


Fig. 1 Southern blot analysis of DNA from cells infected with either STLV-3_{agm} or HTLV-4. Total cellular DNA from HTLV-4-infected HUT-78 cells (lanes 2-5), STLV-3-infected HUT-78 cells (lanes 6-11) and uninfected HUT-78 cells (lane 1) cut with *Bam*HI and *Pvu*II and hybridized with the STLV-3_{agm}-derived probe K2-10-BA. Three HTLV-4-infected cell lines were established by co-cultivation with peripheral blood lymphocytes (PBL) of HTLV-4-infected West African individuals (lane 2, P82; lane 3, P190; lane 4, PK289)¹. The fourth HTLV-4-positive cell line (lane 5, 78(BK-28)) was established by transfection of HUT-78 cells with λ BK-28 DNA (see text). The STLV-3-infected cell lines (lane 6, Mm6324; lane 7, K2; lane 8, K3; lane 9, K6; lane 10, K8; lane 11, K1) were established by co-cultivation of PBL from African green monkeys with HUT-78 cells, except for Mm6324 which was derived by infection of HUT-78 cells with STLV-3_{agm} but subsequently passed *in vivo* in a macaque (ref. 5 and P. J. Kanki *et al.*, manuscript in preparation).

Methods. Total cellular DNA was purified from HUT-78 cells, from six STLV-3_{agm}-infected HUT-78 cell lines and from four HTLV-4-infected HUT-78 cell lines as described⁴⁰. Aliquots (20 μ g) of genomic DNA were digested with *Pvu*II and *Bam*HI, separated by electrophoresis on 1% agarose gels, transferred to nitrocellulose filters, and hybridized for 24 h in the presence of 10% dextran sulphate and 50% formamide at 42 °C as previously described⁴¹. Viral DNA insert from plasmid pK2-10-BA¹¹ was used as probe as described previously^{11,41}.

the presence of an additional site in the envelope region. Significantly, one of these isolates (239) was derived by *in vivo* passage of macaque isolate 251, which did not show the polymorphism. Therefore, the *Pst*I site polymorphism probably represents the emergence of a minor population of virus present in isolate 251, rather than mutation of virus in the animal that gave rise to isolate 239.

Minor variations, corresponding to variant genomes or host-virus junction bands, distinguished each of the cell lines under study and establish that each corresponds to an independently infected cell culture (data not shown). The existence of minor variants within cell lines was confirmed in two cases by the demonstration of restriction sites that differed between independent molecular clones of STLV-3_{agm} (ref. 11) and HTLV-4 (Fig. 2). Nevertheless, preliminary DNA sequence comparisons covering ~0.5 kilobases (kb) of the envelope gene and 1 kb of the 3' *orf*-LTR regions indicate that these regions of STLV-3_{agm} and HTLV-4 are ~99% homologous (see Fig. 3).

Strong homology was also detected between our probes and viral genomes present in cell lines infected with STLV-3_{smm} from the California Regional Primate Research Center¹² (provided by P. Marx) and another macaque isolate, probably derived by cross-species infection with a sooty mangabey virus, from the Delta Regional Primate Research Center (STLV-3_{delta})⁶ (provided by M. Murphey-Corb). The patterns of viral fragments observed in these cells were distinct from each other and from the STLV-3/HTLV-4 consensus (data not shown).

The strong degree of restriction site conservation noted here between multiple virus isolates from three species is largely without precedent in the study of retroviruses. Our results were

* To whom correspondence should be addressed.

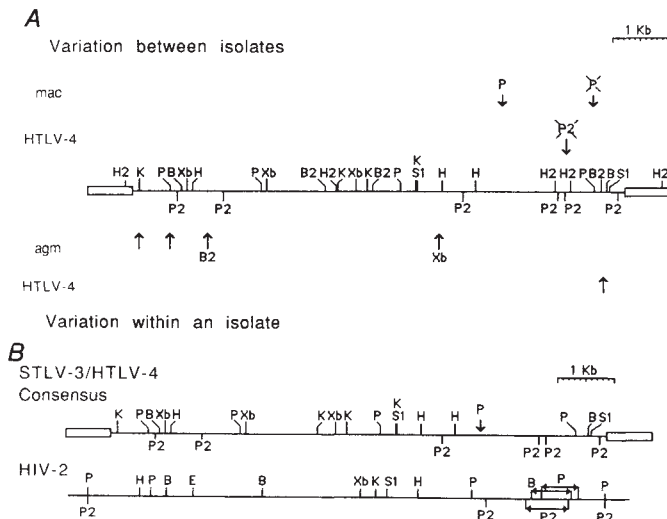


Fig. 2 Restriction site comparison of STLV-3 from African green monkeys and macaques, HTLV-4 and HIV-2. **A**, The consensus map shown corresponds to the major form of viral DNA present in each of six isolates of STLV-3_{agm}, two isolates of HTLV-4, and STLV-3_{mac} isolate 251. Arrows above the map, polymorphic restriction sites observed between isolates. Top line, polymorphisms detected in the major form of viral DNA in cells infected with the three macaque isolates. Sites for *Xba*I were not mapped in any of these isolates and the *Sac*I sites were mapped only in isolate 142. The *Pst*I site unique to the latter two isolates is indicated and the *Pst*I site found in the consensus but not found in these two isolates is indicated with a X drawn through the arrow. The second line indicates the polymorphic sites found in HTLV-4. The crossed-out *Pvu*II site is not found in the major form of viral DNA present in PK-289 cellular DNA or in four molecular clones derived from that cell line. Polymorphic restriction sites observed in molecular clones derived from a single cell line are shown below the map. The first lower line indicates polymorphic restriction sites found in five partial proviral DNA clones derived from one STLV-3_{agm} isolate¹¹. The second lower line indicates the single polymorphic site present in one of four complete proviral HTLV-4 clones. **B**, Alignment of restriction site maps of the STLV-3/HTLV-4 consensus with HIV-2²³. The maps were aligned according to the sequence homology shown in Fig. 3. The position of the *Bam*HI, *Pst*I and *Pvu*II sites near the 3' end of HIV-2 were not consistently defined in ref. 3 and are indicated by horizontal arrows.

Methods. Analyses were performed as described in Fig. 1 legend and in ref. 11, using STLV-3 probes encompassing the entire viral genome. One macaque isolate (251) and all the African green monkey and HTLV-4 isolates were provided under code by P. Kanki. DNA from cells infected with all three macaque isolates were provided uncoded by R. Desrosiers. Except for the PK-289 cell line, each sample was provided on at least one occasion as cell pellets or lysates for direct DNA analysis. Restriction site abbreviations: H2, *Hinc*II; K, *Kpn*I; P, *Pst*I; B, *Bam*HI; B2, *Bgl*II; P2, *Pvu*II; Xb, *Xba*I; H, *Hind*III; S1 *Sac*I. Molecular cloning of HTLV-4 proviral DNA was carried out with *Eco*RI-digested PK-289 DNA and λ J1⁴⁰ vector arms by our standard procedures using K2-2-XbaB¹¹ as probe.

particularly unexpected because HIV-1, the closest known relative of STLV-3/HTLV-4, is, by contrast, one of the most genetically unstable retroviruses studied¹³. It may be that a link exists between the low mutation rate of STLV-3/HTLV-4 and their relative lack of pathogenicity in their natural hosts^{1,5} and *in vitro* (discussed below).

Our data suggest the possibility that STLV-3_{agm}, STLV-3_{mac} and HTLV-4 may represent a remarkably stable virus capable of transmission between several species under conditions of natural exposure, with or without subtle genetic alterations not detected by our analysis. In support of this hypothesis is the finding that STLV-3 infection of macaques has only been repor-

ted in animals that may have been exposed to African green monkeys or mangabeys during captivity^{6,8,9}, suggesting that inter-species transmission under these circumstances does indeed occur. That the viruses under study display remarkable stability is supported by the finding that STLV-3_{mac} isolates 142 and 251 were obtained three years apart and from different macaques yet are identical in restriction-site pattern. Also, isolate 239 was obtained after one year and three sequential *in vivo* passages⁹ yet differs from the other two viruses in only 0-2 of the 23 sites studied.

Transmission of animal retroviruses between species has been clearly shown to occur in other systems under experimental conditions. HTLV-1 infects certain monkeys and rabbits^{14,15} and is 90% homologous throughout the 3' 3 kb of its genome to STLV-1 (ref. 16) which is found to infect naturally as many as 34 species of Old World monkeys¹⁷. Other examples include HIV infection of chimpanzees¹⁸, and murine leukaemia virus infection of rats^{19,20}. Bovine leukaemia virus (BLV) infects chimpanzees²¹ and sheep^{21,22} under experimental conditions and may also infect sheep under natural conditions²³.

An alternative explanation of the origin of these viruses is that some of the STLV-3_{agm}, STLV-3_{mac}, and/or HTLV-4 isolates obtained to date do not correspond to independent virus isolates, but originate from other virus-producing cell cultures maintained in the same laboratories. Consistent with this hypothesis is the finding that STLV-3 isolates made in other laboratories (one from a sooty mangabey and one from a macaque possibly infected from a sooty mangabey^{6,12}) are clearly distinct from the macaque, African green monkey and HTLV-4 isolates we have studied in detail. However, as noted above, each cell line was distinguishable by minor variant or host-virus junction bands (confirming them to be independently infected cultures). Furthermore, the polymorphism of restriction sites detected between and within isolates from one species was not found in isolates from other species.

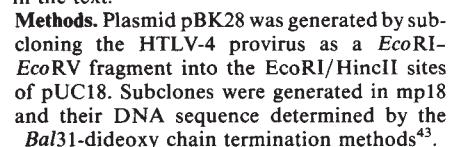
Comparison of the STLV-3/HTLV-4 consensus map with that reported for HIV-2 (ref. 3) reveals that the positions of approximately one-third of the sites are conserved (Fig. 2B). DNA sequence comparison between a region of the LTR of HIV-2 and the corresponding region of STLV-3 (V.H., N.R. and J.I.M., manuscript submitted) and HTLV-4 reveals that the sequences are largely colinear and share stronger homology than either sequence shares with HIV-1 in the regions examined (Fig. 3A). STLV-3/HTLV-4 and HIV-2 share 93% homology in the R region, whereas several gaps must be introduced to maintain alignment of either sequence with that of HIV-1 (Fig. 3A). In regions outside the gaps required for alignment HIV-2 shares slightly greater homology with HIV-1 than does STLV-3/HTLV-4 with HIV-1, especially in the Sp1 binding region (see below).

In the portion of the U3 region examined, introduction of a gap is required to maintain alignment between STLV-3/HTLV-4 and HIV-2 sequences and aside from this gap they share 72% homology. Interestingly, each virus conserves the 5'-GGGACTTTC-3' motif (boxed in Fig. 3B) found in the enhancer region of simian virus 40 (SV40)²⁴, in the κ 3 light chain immunoglobulin enhancer²⁵, and perfectly conserved (including a second copy upstream from this position) in all HIV-1 strains sequenced to date. This corresponds to the longest stretch of perfect homology between the STLV-3/HTLV-4 and HIV LTRs (V.H., N.R. and J.I.M., submitted), and is immediately upstream of several potential Sp1 transcription factor binding sequences²⁶ (individually boxed in Fig. 3B). HIV-2 contains four copies of a sequence similar to the Sp1 consensus (5'-G/TGGGC/AGGPuPy-3') whereas STLV-3/HTLV-4 and HIV-1 contain three copies. The Sp1 sequences found in HIV_{SF2} have been shown to bind and be transcriptionally responsive to Sp1²⁶.

The sequence responsive to transactivation in HIV-1 (*tar*) is found in the R region of the LTR²⁷ within a large inverted repeat²⁸. A distinct inverted repeat is present within the STLV-

All four cloned HTLV-4 proviruses (BK-27, 28, 44, 48) were introduced into HUT-78 cells by transfection^{29,30} and cultures monitored for cytopathic effects (CPE) and the presence of Mg²⁺-dependent reverse transcriptase (RT) activity as described for HIV-1³¹. Cellular atypia was first detected in BK-28-transfected cells at day 5 and numerous multinucleated giant cells were observed by day 10 after transfection (Fig. 4A). Reverse transcriptase activity in these cells rose to 7.2×10^5 c.p.m. ml⁻¹ at 10 days, compared with levels of $1.3\text{--}2.9 \times 10^3$ c.p.m. ml⁻¹ in mock-transfected controls and BK-27-, BK-44-, and BK-48-transfected cells. A similar transition of CPE and RT activity was observed in the BK-44-transfected cells at around day 20 after transfection. RT activity has remained elevated in the BK-44- and BK-28-transfected cells for more than 90 days. In contrast, morphological changes diminished progressively and by 30 days only few giant cells were present (Fig. 4A).

Production of biologically active viruses was confirmed by infection of HUT-78 cells with filtered 78(BK-28) cell culture supernatant. Cellular atypia, giant cells, and increased RT activity were noted at day 6 after infection but not in mock-infected HUT-78 controls. As with transfected cells, RT activity persisted at high levels, minimal cytolysis was observed, cultures survived without the addition of fresh HUT-78 cells and the presence of viral sequences was confirmed by Southern blot analysis (data not shown).



The T-cell surface protein CD4 is a receptor for HIV-1³²⁻³⁵ and infection can be inhibited by monoclonal antibodies to CD4. To determine the role of CD4 in HTLV-4 infection, we incubated HUT-78 cells with OKT4A, a monoclonal antibody to CD4, or monoclonal anti-HLA-Dr, which recognizes a separate cell surface protein on these cells³⁶. Cells were then exposed to filtered 78(BK-28) supernatant (multiplicity of infection >0.5

tissue culture infectious units) for 30 minutes, washed extensively, and cultured in the continuous presence of OKT4A or HLA-Dr antibodies. On day 8 after infection, CPE and a 70-fold rise in RT activity were observed in HLA-Dr-treated cells but not in cells maintained in the presence of OKT4A (data not shown). Thus, our results indicate that the CD4 protein serves as a component of the receptor for HTLV-4 on HUT-78 cells.

Binding of CD4⁺ uninfected cells to viral envelope proteins expressed on the surface of HIV-1-infected cells has been proposed as a mechanism for the induction of syncytium formation and cell killing by HIV-1³⁷⁻³⁹. Although infection with cloned HTLV-4 results in giant cell formation, there is no evidence for significant cell killing. Comparative analysis of STLV-3/HTLV-4, HIV-2 and HIV-1 may provide a better understanding of the precise mechanism of cytolysis by the latter. Furthermore, as the structural and genetic relationship between HTLV-4, the STLV-3s, HIV-2 and HIV-1 become more precisely defined, it may be possible to determine specifically what differences are responsible for the lack of stability, receptor interactions and pathogenesis of HIV-1.

We thank M. Essex, P. Kanki, R. C. Desrosiers, M. Murphey-Corb, and P. Marx for virus infected cells and cell DNAs; G. Goldstein for the monoclonal antibodies; P. Kanki for performing the radioimmunoprecipitation analysis; and M. Essex and P. Kanki for comments on the manuscript. This work was supported by the Massachusetts AIDS Research Council and by NIH grants to J.I.M. H.K. is recipient of a Clinical Investigator Award from the NIH, N.R. is a fellow of the Deutsche Forschungsgemeinschaft, and V.H. is a fellow of the MRC of Canada. The HTLV-4 sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00269.

Received 2 December 1986; accepted 3 March 1987.

1. Kanki, P. J. *et al. Science* **232**, 238-243 (1986).
2. Clavel, F. *et al. Science* **233**, 343-346 (1986).
3. Clavel, F. *et al. Nature* **324**, 691-695 (1986).
4. Kanki, P. J. *et al. Lancet* **i**, 1330-1332 (1985).
5. Kanki, P. J., Alroy, J. & Essex, M. *Science* **230**, 951-954 (1985).
6. Murphey-Corb, M. *et al. Nature* **321**, 435-437 (1986).
7. Fultz, P. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5286-5290 (1986).
8. Kanki, P. J. *et al. Science* **228**, 1199-1201 (1985).
9. Daniel, M. D. *et al. Science* **228**, 1201-1204 (1985).
10. Benveniste, R. E. *et al. J. Virol.* **60**, 483-490 (1986).
11. Hirsch, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9754-9758 (1986).
12. Lowenstein, L. J. *et al. Int. J. Cancer* **38**, 563-574 (1986).
13. Coffin, J. M. *Cell* **46**, 1-4 (1986).
14. Yamamoto, N. *et al. Med. Microbiol. Immun.* **173**, 57-64 (1984).
15. Miyoshi, I. *et al. in Retroviruses in Human Lymphoma/Leukemia* (eds Masano, M., Sugano, H., Sugimura, T. & Weiss, R. A.) 121-127 (Japan Sci. Soc. Press/VNU Science Press, Utrecht, 1985).
16. Watanabe, T. *et al. Virology* **144**, 59-65 (1985).
17. Hayami, M. *et al. Int. J. Cancer* **33**, 179-183 (1984).
18. Alter, J. J. *et al. Science* **226**, 549-552 (1984).
19. Graffi, A. & Gimmy, J. *Naturwissenschaften* **44**, 518-519 (1957).
20. Gross, L. *Proc. Soc. exp. Biol. Med.* **112**, 939-945 (1963).
21. McClure, H. M. *et al. Cancer Res.* **34**, 2745-2757 (1974).
22. Mammerickx, M., Dekegel, D., Burny, A. & Portetelle, D. *Vet. Microbiol.* **1**, 347-350 (1976).
23. Onuma, M., Baumgartner, L. E., Olson, C. & Pearson, L. D. *Cancer Res.* **37**, 4075-4081 (1977).
24. Zenke, M. *et al. EMBO J.* **5**, 387-397 (1986).
25. Sen, R. & Baltimore, D. *Cell* **46**, 705-716 (1986).
26. Jones, K. A., Kadonaga, J. T., Luciw, P. A. & Tijan, R. *Science* **232**, 755-759 (1986).
27. Rosen, C. A., Sodroski, J. G. & Haseltine, W. A. *Cell* **41**, 813-823 (1986).
28. Okamoto, T. & Wong-Staal, F. *Cell* **47**, 29-35 (1986).
29. Milman, G. & Herzberg, M. *Somatic Cell Genet.* **7**, 161-170 (1981).
30. Dorsett, D. L., Ilana, K. & Winocour, E. *J. Virol.* **48**, 218-228 (1983).
31. Popovic, M., Sargadharan, M. G., Read, E. & Gallo, R. C. *Science* **224**, 497-500 (1984).
32. Dagleish, A. G. *et al. Nature* **312**, 763-767 (1984).
33. Klatzmann, D. *et al. Nature* **312**, 767-768 (1984).
34. McDougal, J. S. *et al. Science* **231**, 382-385 (1986).
35. Maddon, P. J. *et al. Cell* **47**, 333-348 (1986).
36. Ceuppens, J. L., Goodwin, J. S. & Searles, R. P. *Cell. Immun.* **64**, 277-292 (1981).
37. Folks, T. *et al. Science* **231**, 600-602 (1986).
38. Lifson, J. D., Reyes, G. R., McGrath, M. S., Stein, B. S. & Engleman, E. G. *Science* **232**, 1123-1127 (1986).
39. Sodroski, J., Goh, W. C., Rosen, C., Campbell, K. & Haseltine, W. A. *Nature* **322**, 470-474 (1986).
40. Mullins, J. I., Brody, D. S., Binari, R. C. Jr & Cotter, S. M. *Nature* **308**, 856-858 (1984).
41. Mullins, J. I., Casey, J. W., Nicolson, M. O. & Davidson, N. *Nucleic Acids Res.* **8**, 3287-3305 (1980).
42. Sanchez-Pescador, R. *et al. Science* **227**, 484-492 (1985).
43. Poncz, M., Solowiejczyk, M., Schwartz, E. & Surrey, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4298-4302 (1982).

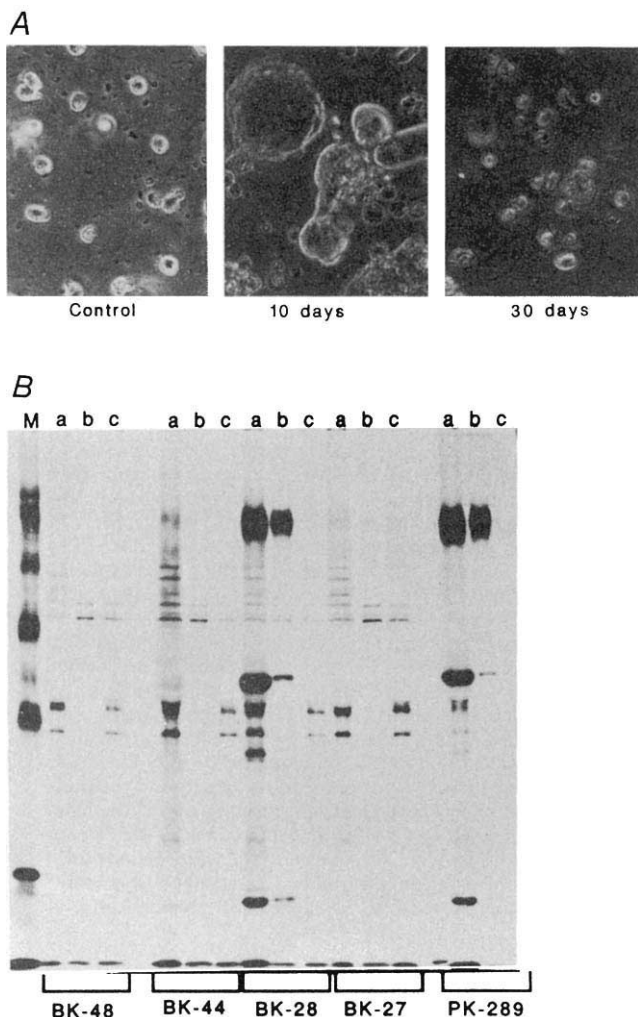


Fig. 4 Biological activity of the HTLV-4 clone BK-28. **A**, Photomicrographs of BK-28-transfected HUT-78 cells at 10 and 30 days and mock-transfected HUT-78 control cells (phase contrast, $\times 10$). **B**, RIP-SDS/PAGE analysis of the PK-289 cell line and of HUT-78 cell cultures transfected with λ clones of BK-48, BK-44, BK-28 and BK-27. Lysates of each cell culture were reacted with the following sera: lanes a and b, sera from two healthy West African individuals seropositive for HTLV-4; lane c, serum from a patient with adult T-cell leukaemia who was seronegative for HTLV-4 and HIV-1.

Methods. Transfection of λ clones was by a modified DEAE-dextran method^{30,31}. HUT-78 cells (6×10^6) were washed in phosphate-buffered saline (PBS) and treated with trypsin (0.05 mg ml^{-1}) in Hank's Buffered Balanced Salts (without Ca^{2+} or Mg^{2+}) plus 5 mM EDTA for 4 min. Cells were then washed in TD buffer (25 mM Tris-HCl pH 7.4, 140 mM NaCl, 5 mM KCl, 0.7 mM K_2HPO_4) and resuspended in 1 ml of TD buffer containing 200 μg DEAE-dextran and 1 μg of DNA. Cells were incubated at room temperature for 20 min and the transfection terminated by the addition of 5 ml RPMI containing 20% fetal calf serum. Radioimmunoprecipitation and SDS-polyacrylamide gel electrophoresis were carried out on day 7 after transfection by published methods¹. Markers for relative molecular mass are, in ascending order, 30, 46, 69, 92.5 and 200 thousand.

A clue to the basic defect in cystic fibrosis from cloning the CF antigen gene

Julia R. Dorin*, Michal Novak†§, Robert E. Hill*,
David J. H. Brock†, David S. Secher†
& Veronica van Heyningen*

* MRC Clinical and Population Cytogenetics Unit, and † Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK
† MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The metabolic basis of the autosomal recessive disease cystic fibrosis (CF) remains unidentified. Elevated levels of a serum protein in CF homozygotes and obligate heterozygotes have been described¹. As heterozygotes are clinically unaffected, any consistently observed abnormality in these individuals is a likely pointer to the aetiology of the disease. The gene for this serum protein, called cystic fibrosis (CF) antigen, has been mapped to chromosome 1 (ref. 2). It is not the gene that is mutant in CF because this has since been assigned to chromosome 7 by cosegregation of the disease with closely linked DNA markers in CF families³⁻⁵. CF antigen is a product of normal and leukaemic granulocytes² and is inducible in the promyelocytic cell line HL60 (M.N., J.D., C. Hayward, F. Northrop, D.J.H.B., J. Walker, V. van H. and D.S.S., manuscript in preparation). We have isolated cDNA clones for this protein from a library constructed with messenger RNA from chronic myeloid leukaemia (CML) cells. The complete nucleotide sequence was obtained from the cDNA clone and by primer extension of mRNA. We have confirmed that the gene encoding CF antigen is on chromosome 1 and have localized it to a particular region. RNA blot analysis shows a 550-bp major transcript in CML cells and in induced HL60. The amino-acid sequence predicted from the nucleotide sequence shows significant homology with intestinal and brain calcium-binding proteins. Abnormal accumulation of such a protein in CF is a clue which must be pursued now that evidence is gathering that the basic defect in CF is in pathways controlling chloride channel activity^{6,7}.

Quantification of serum levels of CF antigen by immunoassay shows that the three genotypes (CF homozygotes, heterozygotes and normals) fall into three overlapping classes¹. Monoclonal antibodies, recognizing two epitopes on this protein, have been produced⁸. These antibodies have now been used for immunopurification and sequence-comparison of the antigen from a variety of sources (M.N. *et al.*, in preparation).

Partial amino-acid sequence analysis of the purified protein allowed us to synthesize a highly redundant oligonucleotide probe, corresponding to the region indicated in Fig. 1 (M.N. *et al.*, in preparation), which was used to isolate the cDNA clone CFA 8-9. The library screened was constructed in λ gt10, using mRNA prepared from CML cells. Antigen turnover had previously been demonstrated in such cells by the specific immunoprecipitation of labelled protein following ³⁵S-methionine incorporation. The identity of CFA 8-9 was confirmed by matching the predicted and known amino-acid sequence outside the oligonucleotide parent region.

The nucleotide sequence of CFA 8-9 is shown in Fig. 1. The insert of 338 nucleotides includes an open reading frame of 250 base pairs (bp) which encodes 83 amino acids stretching from the first amino acid used for the construction of the oligonucleotide probe, to the C-terminal residue before the chain-termination codon. The sequence encoding the first 11 N-terminal amino acids, known from the protein sequence data, is also shown in Fig. 1. The nucleotide sequence for this and the 5'-untranslated region (Fig. 1) was determined by oligonucleotide primer-extension

```
Sense strand:
-51
CTCTTGTCAGCTGTCTTTCAGAAGACCTGGTGGGNAAGTCGTCGGGCATC

1                               G*****          60
ATGTTGACCGAGCTGGAGAAAGCCTTGAACCTATATCGACGCTTACCACAAGTACTCC
[M L T E L E K A L N S I I D V Y H K Y S

                               CFA 8-9
120
CTGATAAAGGGAATTTCCATGCCGTCTACAGGGATGACCTGAAGAAATTGCTAGAGACC
L I K G N F H A V Y R D D L K K L L E T

180
GAGTGTCTCTAGTATATCAGGAAAAAGGGTGCAGACGCTGGTTCAAAGAGTTGGATATC
E C P Q Y I R K K G A D V I W F K E L D I

240
AACACTGATGGTGCAGTTAACTTCCAGGAGTTCCTCATTCTGGTGATAAAGATGGCGTGG
N T D G A V N F Q E F L I L V I K M A W

300
CAGCCCCACAAAAAGCCATGAAGAAAGCCACAAAGAGTAGCTGAGTTACTGCCAGAGG
Q P T K K A M K K A T K S S !

CTGGGCCCTGACATGTACCTGCAGAATAATAAGTCATCAATACCTC (polyA)
```

Fig. 1 Nucleotide and corresponding amino-acid sequence of the cDNA clone CFA 8-9. The probable polyadenylation signal of the mRNA is underlined. Asterisks, the oligonucleotide used to isolate the clone; !, termination codon; vertical bar, start of probe; brackets, N-terminal amino acids determined by protein sequencing (M.N. *et al.*, in preparation). The nucleotide sequence from -51 to +33 was determined by reverse transcriptase primer extension using induced HL-60 cell mRNA as a template.

Methods. RNA was prepared³⁰ from 10¹¹ CML cells collected by leukopheresis. Poly(A)⁺ RNA was isolated³¹ and used as a template for cDNA synthesis³². The resulting double-stranded cDNA was cloned in λ gt10 by standard procedures³³. Titration on *Escherichia coli* Hfl strain revealed a library of 70,000 independent recombinants. The amplified library was screened by the method of Benton and Davis³⁴. The 144-fold redundant 15-mer oligonucleotide probe was labelled with polynucleotide kinase using 10 pmol of 5' ends and 200 μ Ci of [γ -³²P]ATP per reaction³⁵. Filters were washed in 2 $\times$ SSC containing 0.1% SDS for 90 min at 37 °C. Positive plaques were re-cloned repeatedly and then sub-cloned into M13mp9 for sequencing. Both strands of the clone with the largest insert (CFA 8-9) were sequenced by the dideoxy chain termination method³⁶. Reverse transcriptase mediated primer extension and dideoxy sequencing of first strand cDNA was carried out following Novak *et al.* (manuscript in preparation).

tion using mRNA from DMSO-induced HL-60 as template. The relative molecular mass (M_r) of 10,938 deduced for the complete sequence of 94 amino acids, is in agreement with the approximate subunit M_r observed for immunoprecipitated antigen (M.N. *et al.*, in preparation). There are no N-glycosylation sites predicted from the amino-acid sequence and there is no evidence that CF antigen is a glycoprotein. The 3' untranslated region of the mRNA is unusually short. On re-screening the CML cDNA library with CFA 8-9 as probe, approximately 0.5% of the recombinants were positive, suggesting that the gene product is relatively abundant in these cells.

RNA blot analysis (Fig. 2) on a variety of tissues and cell lines was carried out to examine message size and abundance. In all cases where mRNA hybridizing to the CFA 8-9 insert was observed, two bands were seen. The major band of approximately 550 bp was of at least 20-fold higher intensity than the second 1 kilobase transcript. Total mRNA levels loaded were monitored by re-probing with an actin cDNA probe⁹. Poly(A)⁺ mRNA isolated from two CML samples, one of which had been used for the construction of the cDNA library, contained very high levels of CFA 8-9 transcripts. Equivalent amounts of mRNA isolated from uninduced HL60 cells gave no detectable signal, whereas HL60 cells induced to differentiate along the myeloid pathway¹⁰ possess both size transcripts. However, retinoic acid (RA) and dimethyl sulphoxide (DMSO) induce

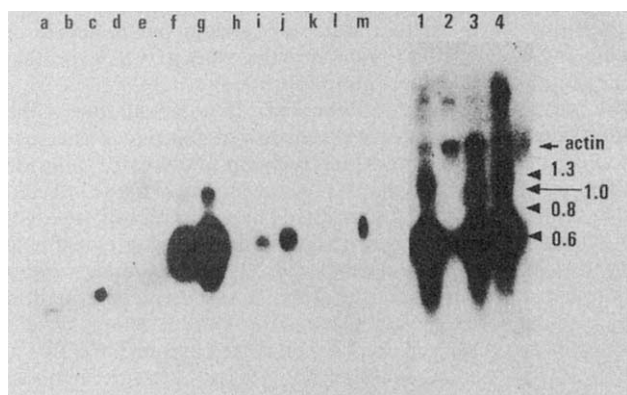
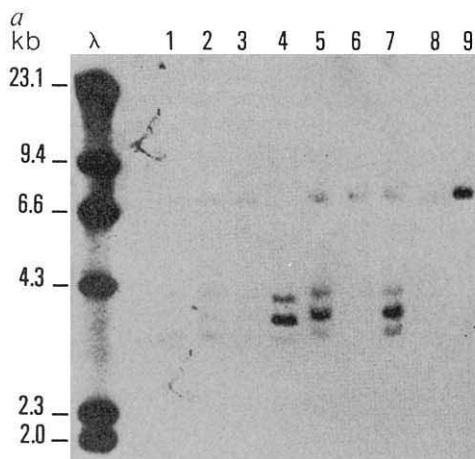


Fig. 2 RNA (Northern) blot analysis of various human cell lines and tissues: a, Hep G2 (hepatoma)³⁷; b, EMBK (embryonic kidney, ATCC No. CRL 1573); c, MNNG-HOS (osteosarcoma, ATCC No. CRL 1547); d, Daudi (Burkitt lymphoma derived B-lymphoblastoid line)³⁸; e, hepatoma; f, CML #1; g, CML #2; h, uninduced HL60; i, HL60 induced (2 days RA); j, normal liver; k, WEHI TG (mouse myeloid stem cell)¹²; l, WEP 1342.2.4 (mouse × human myeloid cell hybrid containing human chromosomes 7, 10, 13, 20)²; m, WEH 3127.1.9 (myeloid cell hybrid containing human chromosomes 1, 9, 14, 21 and expressing CFAg)². Lane 1, CML #1; 2, HL60 induced (2 days RA); 3, HL60 induced (5 days DMSO); 4, CML #2. Both blots were probed with CF antigen gene cDNA, and reprobed with mouse actin to ensure comparable sample loading (reprobe shown for lanes 1–4 only).

Methods. Cells were grown in RPMI 1640 medium supplemented with 5% fetal calf serum, except for uninduced HL60 which was grown in serum-free medium supplemented with insulin, transferrin and selenite (ITS, Sigma). HL60 induction was by continuous culture with 1.2% (v/v) DMSO or 10⁻⁶ M RA. RNA was isolated as in Fig. 1. Then poly(A)⁺ RNA (2 µg in lanes a, b, c, k, l, m), total RNA (20 µg in lanes d–j) or poly(A)⁺ RNA (1 µg in lanes 1–4) was separated on a 1.2% agarose/formaldehyde gel and transferred directly to nitrocellulose³⁵. Hybridization was at 68 °C with the insert from CFA 8-9 randomly primed with a mixed sequence hexadeoxynucleotide (Pharmacia), filled in using Klenow enzyme and radiolabelled with [α -³²P]dTTP. Filters were washed in 2 × SSC/0.1% SDS at 68 °C for 90 min.

low and high relative message abundance respectively. A possible explanation for this may be that cells treated with RA and DMSO do not reach the same degree of maturation along the myeloid pathway. After equivalent treatment regimes only 10% of DMSO-treated cells are recognizable as mature neutrophils^{10,11}, compared with 30% for RA treatment. Messenger RNA isolated from a variety of other malignant human cell lines (embryonic kidney, osteosarcoma, hepatoma and Burkitt's lymphoma) gave no detectable signal. Human liver mRNA gave a weak signal which we have assumed to be due to blood contamination of the original tissue, because we have demonstrated antigen turnover in normal granulocytes, and immunohistochemical analysis of liver sections reveals no antigen in hepatocytes (V. van H. *et al.*, unpublished). No detectable signal was seen with mRNA from the mouse myeloid stem cell line WEHI-TG¹², or in somatic cell hybrids between WEHI-TG and human CML cells unless human chromosome 1 was present in such hybrids². It appears therefore that there is no synthesis of a homologous cross-hybridizing mRNA in this mouse myeloid line. No signal distinguishable from the human 550-bp band was found in the 'switched on' chromosome 1 containing hybrids. Thorough immunohistochemical analysis of normal tissues revealed CF antigen only in a few non-keratinizing squamous epithelia (tongue and oesophagus), in addition to granulocytes (manuscript in preparation).

Human and mouse DNA digested with restriction enzymes *Pst*I, *Bam*HI, *Eco*RI and *Hind*III give a single hybridizing band with CFA 8-9 insert even with low stringency washes. Digestion with *Bgl*II gives three bands but of small size, consistent with restriction sites within the gene. This suggests the existence of only one gene for this protein in both these species, and the absence of any closely related cross-hybridizing genes. Mouse-human somatic cell hybrids, regardless of whether they express CF antigen², exhibit a band of human mobility only when chromosome 1 is present at high frequency (Fig. 3a). The confirmation that this gene is on chromosome 1 is additional evidence that the cloned insert corresponds to the immunologically recognized protein^{1,2}. The availability of cell hybrids containing defined fragments of chromosome 1 has allowed us to sublocalize the gene locus to the region between bands q12 and q22 on the long arm (Fig. 3b).



b

Human chromosomes present	Chromosomal localization of the gene for CF antigen																											
	Hybrid	Fragment of 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	CFAg		
Rag J8	pter-q21																										+	
SK82	pter-q12																		q								-	
SK81	pter-p31																										-	
NYX 5.6	complete	not fully characterized																										+
NYX 5.6.11	pter-q23	not fully characterized																										+
F9.6	complete																										+	
WEH 3127.1.9 T ₇	complete																										+	
WEH 3127.1.9 T ₈	absent																										-	

Fig. 3 a, Southern blot analysis³⁹ to illustrate mapping of CFA 8-9 in rodent/human somatic cell hybrids. Lane 1, NYX5.6.11⁴⁰; 2, F9.6 (unpublished data); 3, NYX5.6⁴⁰; 4, CML #3; 5, WEH 3127.1.9 T7²; 6, WEH 3127.1.9 T8; 7, WEH 3127.1.9 T7; 8, WEH 3127.1.9 T8; 9, WEHI-TG. Size markers (λ fragments) in kb. b, Summary of chromosome analysis of the location of the gene for CF antigen (CFAg). Segregation of the CFA 8-9 hybridizing, human-specific restriction fragment with chromosome 1 or fragments of it. WEH 3127.1.9 T7 and T8 are subclones of the same hybrid², FACS sorted for expression and non-expression respectively of a chromosome-1-encoded cell surface marker⁴¹. Black squares denote the presence of chromosomes in >50% of cells scored; q, long arm only present. The restriction fragments corresponding to the human gene for CFA 8-9 (+ in CFAg column) are seen whenever the entire chromosome 1 is present and also in hybrid RAG J8 with human chromosome 1 from the tip of the short arm to band q21 on the long arm (pter-q21) present. No human CFA8-9 bands are present (- in CFAg column) in hybrid SK 82 which contains only region pter-q12 of chromosome 1. The gene must therefore lie between bands q12 and q22 on chromosome 1. Other hybrid cell data support this result.

Methods. In a, ~10 µg of genomic DNA was digested with *Bgl*II, fractionated on 0.8% agarose gels and transferred to nitrocellulose filters. After hybridization with labelled cDNA probe CFA 8-9 insert, blots were washed in 2 × SSC/0.1% SDS at 68 °C.



Fig. 4 Alignment of the predicted amino-acid sequence of CF antigen with amino-acid sequences of ICaBP from cow, pig and rat; bovine S100a and b subunits; porcine p11 and human 2A9. The calcium binding ligands, as determined by X-ray crystallography²⁰ for the EF-hand region of bovine ICaBP, are marked with asterisks. Calcium-binding ligands marked with triangles are from the second distinct type of calcium-binding region in ICaBPs and S100. Amino-acid residues which are identical in CF antigen and at least one of the other sequences are boxed.

The deduced amino-acid sequence of the cloned CF antigen gene was compared with all sequences in the NBRF database by a method which permits high stringency comparison with all the sequences present¹³. Highly significant homology (optimized score: 189, 43% identity over a stretch of 72 amino acids) was obtained with bovine S100a subunit, a brain-associated calcium-binding protein¹⁴, and the similarity extends to the homologous S100b subunit. Strong homology was also seen with bovine intestinal calcium-binding protein (BICaBP)¹⁵ and its pig and rat homologues. These proteins all belong to a class of calcium-binding proteins which are postulated to be the major transducers of biological calcium signals¹⁶. Like S100 proteins, but unlike the intestinal calcium-binding proteins which are monomeric, native cystic fibrosis antigen is predominantly in dimeric form under non-reducing conditions (M.N. *et al.*, manuscript in preparation). Amino-acid comparisons are shown in Fig. 4. Two additional sequences, not yet in the NBRF database but recently shown to be similar to S100 proteins, are included. These are: the sequence for porcine p11 (or p10), a regulatory subunit of the protein complex which is a major cellular target for tyrosine kinase^{17,18}; and the deduced sequence for a cell-cycle specific human cDNA molecule, 2A9¹⁹, the protein product of which has not yet been studied.

In the well studied calcium-binding proteins there are two types of calcium-binding sites²⁰. Amino acids which act as ion-binding ligands in the conventional EF hand-like site²⁰ are marked with asterisks in Fig. 4. These are well conserved in cystic fibrosis antigen, suggesting that calcium binding remains a normal function of this protein. The criteria for the conservation of calcium-binding capacity at the second site (marked with triangles) are less clearly defined²⁰. Equilibrium binding studies, using ⁴⁵Ca²⁺ in competition with unlabelled CA²⁺, have shown

that CF antigen binds two atoms of calcium per molecule of protein and with an affinity similar to that shown by S100 protein (D.J.N.B. and C. Hayward, unpublished data).

The most widely accepted abnormality demonstrable at the cellular level in CF is defective transport of chloride ions across epithelia²¹. However, recent patch-clamp analyses of chloride channel activity in epithelial cells of the airway, a tissue affected in CF, suggest that the conductance characteristics of isolated, cell-free chloride channels from CF tissue are indistinguishable from those observed for normal cells^{6,7}. When similar studies are carried out with whole cells, the β -adrenergic stimulation of chloride channel activity observed in normal cells is absent in CF cells. As cAMP generation is unaltered in such CF cells^{6,22}, deviation from normal function might be sought further on in the cAMP pathway or in other signal transducing systems. Activation of the phosphatidylinositol pathway²³ leads to mobilization of intracellular calcium. Alteration of any component of this system is a candidate for the basic defect in CF tissue. Recent investigations have suggested the existence of more than one receptor, capable of responding through different pathways to signals by certain ligands^{24,25}. Control of chloride channel activity may fall into this category.

The most likely tissue source for the elevated serum levels of cystic fibrosis antigen is granulocytes. This would imply that myeloid cells are an affected tissue in CF. There have been suggestions that this is the case^{26,27}, although no clinically relevant granulocyte abnormalities have been reported²⁸. Investigation of ion transport with particular reference to calcium metabolism in granulocytes might be fruitful. Calcium-modulated activation of ion channels has been demonstrated in normal human neutrophils²⁹. Confirmation that the basic defect is expressed in granulocytes would be an important step forward for CF research because these cells are relatively easily accessible from affected and heterozygous individuals.

Demonstration of granulocytes as a tissue affected in CF would suggest that elevated serum levels of cystic fibrosis antigen can be attributed to the absence (in homozygotes) or reduction (in heterozygotes) in these cells of the normal product of the CF gene on chromosome 7 which may be necessary for correct assembly of a multi-molecular complex.

We thank Dr Andrew Coulson for running the database comparison, Dr John Goldman for supplying leukaphoresis samples, Richard Buckland for supplying some hybrid-cell DNA, Judy Fletcher for the chromosome analyses, Richard Meehan and Nick Hastie for invaluable advice and discussion and Professor H. J. Evans for support. J.R.D. is the recipient of an MRC Training Award. M.N. is a Royal Society Exchange Fellow from the Slovak Academy of Sciences, Bratislava. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00278.

Received 25 November 1986; accepted 11 February 1987.

- Bullock, S., Hayward, C., Manson, J. C., Brock, D. J. H. & Raeburn, J. A. *Clin. Genet.* **21**, 336-341 (1982).
- Van Heyningen, V., Hayward, C., Fletcher, J. & McAuley, C. *Nature* **315**, 513-515 (1985).
- Knowlton, R. G. *et al.* *Nature* **318**, 380-382 (1985).
- White, R. *et al.* *Nature* **318**, 382-384 (1985).
- Wainwright, B. J. *et al.* *Nature* **318**, 384-385 (1985).
- Welsh, M. J. & Liedtke, C. M. *Nature* **322**, 467-470 (1986).
- Frizzel, R. A., Rechtemer, G. & Shoemaker, R. L. *Science* **233**, 558-560 (1986).
- Hayward, C., Chung, S., Brock, D. J. H. & Van Heyningen, V. *J. Immun. Meth.* **91**, 117-122 (1986).
- Minty, A. J. *et al.* *J. Biol. Chem.* **256**, 1008-1014 (1981).
- Breitman, T. R., Scolnick, S. E. & Collins, S. J. *Proc. natn. Acad. Sci. U.S.A.* **277**, 2936-2940 (1980).
- Ferrero, D., Pessando, S., Pagliardi, G. & Rovera, G. *Blood* **61**, 171-179 (1983).
- Geurts van Kessel, A. H. M., Tetteroo, P. A. T., von dem Borne, A. E. G., Hagemeijer, A. & Bootsma, D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3748-3752 (1983).
- Lyall, A., Hill, C., Collins, J. F. & Coulson, A. F. W. in *Parallel Computing '85* (eds Feilomeier, M., Joubert, G. & Scheudel, U.) 235-240 (North-Holland/Elsevier, Berlin, 1986).
- Isobe, T. & Okuyama, T. *Eur. J. Biochem.* **116**, 79-86 (1981).
- Fullmer, C. S. & Wasserman, R. H. *J. Biol. Chem.* **256**, 5669-5674 (1981).
- Van Eldrick, L. J., Zendegui, J. G., Marshak, D. R. & Watterson, D. M. *Int. Rev. Cytol.* **77**, 1-61 (1982).

17. Gerke, V. & Weber, K. *EMBO J.* **4**, 2917–2920 (1985).
18. Glenney, J. R. & Tack, B. F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7884–7888 (1985).
19. Calabretta, B., Battini, R., Kaczmarek, L., de Riel, J. K. & Baserga, R. *J. biol. Chem.* **261**, 12628–12632 (1986).
20. Kretsinger, R. H. *Ann. N.Y. Acad. Sci.* **356**, 14–19 (1980).
21. Case, M. *Nature* **322**, 407 (1986).
22. Sato, K. & Sato, F. *J. clin. Invest.* **73**, 1763–1771 (1984).
23. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
24. Michell, R. H. & Houslay, M. *Trends biochem. Sci.* **11**, 239–241 (1986).
25. Wakelam, M. J. O., Murphy, G. J., Hruby, V. J. & Houslay, M. *Nature* **323**, 68–71 (1986).
26. Galant, S. P., Norton, L., Herbst, J. & Wood, C. J. *clin. Invest.* **68**, 253–258 (1981).
27. Davis, P. B., Dieckman, L., Boat, T. F., Stern, R. C. & Doershuk, C. F. *J. clin. Invest.* **71**, 1787–1795 (1983).
28. Talamo, R. C., Rosenstein, B. J. & Berninger, R. W. in *The Metabolic Basis of Inherited Disease* (eds Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L. & Brown, M. S.) 1889–1917 (McGraw-Hill, New York, 1983).
29. Von Tscherner, V., Prod'homme, B., Baggiolini, M. & Reuter, H. *Nature* **324**, 369–372 (1986).
30. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
31. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
32. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263–269 (1983).
33. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A Practical Approach* Vol 1 (ed. Glover, D.) 49–78 (IRL Press, Oxford, 1985).
34. Benton, W. & Davis, R. *Science* **196**, 180–182 (1977).
35. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
36. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
37. Aden, D. P., Fogel, A., Damjanov, I., Plotkin, B. & Knowles, B. *Nature* **282**, 615 (1979).
38. Klein, E. *et al. Cancer Res.* **28**, 1300–1310 (1968).
39. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
40. Van Heyningen V. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8592–8596 (1985).
41. Andrews, P. W. *et al. Ann. hum. Genet.* **49**, 31–39 (1985).

Amino-acid sequence of glycoproteins encoded by three alleles of the *S* locus of *Brassica oleracea*

J. B. Nasrallah*, T.-H. Kao†‡, C.-H. Chen*, M. L. Goldberg‡ & M. E. Nasrallah*

* Sections of Plant Biology, † Biochemistry, Molecular and Cell Biology, and ‡ Genetics and Development, Cornell University, Ithaca, New York 14853, USA

The pollen–stigma interaction of self-incompatibility in the mustard family Brassicaceae is controlled by the *S* locus. Nearly fifty allelic variants at this locus control the specificity of this interaction. Identity of alleles in pollen and stigma results in an incompatible reaction and the arrest of pollen development. Genetic and developmental analyses of self-incompatibility in *Brassica* have identified a class of stigma glycoproteins exhibiting allele-specific isoforms as being among the products of the *S*-locus^{1–3}. A cDNA clone encoding one of these *S*-locus specific glycoproteins has been isolated from *B. oleracea*⁴, and recently, N-terminal protein sequence analysis has shown extensive similarity between three *S*-locus specific glycoproteins from *B. campestris*⁵. Here, we describe cDNA clones encoding such glycoproteins from three different *S*-allele homozygous genotypes, and report the complete amino-acid sequence of the glycoproteins based on the nucleotide sequence of the cDNA clones. We identify allele-encoded differences in the protein sequences which may be relevant to the allelic specificity of the stigma component of the pollen–stigma interaction. The sequences from the sporophytically determined incompatibility system of *Brassica* show no similarity to the reported *S*-locus-associated cDNA sequence of the gametophytic *Nicotiana* self-incompatibility type⁶.

Clones (of λ gt10) containing complementary DNA inserts encoding *S*-locus-specific glycoproteins (SLSGs) were selected by their homology to the SLSG-cDNA previously cloned in pBR322 and designated pBOS5⁷. These λ gt10 clones were derived from three *B. oleracea* lines homozygous for the *S*₆, *S*₁₃ and *S*₁₄ alleles, and extended further at the 5'-end than the original 1,284-base pair (bp) pBR322 clone.

The complete nucleotide sequences were determined for both strands of two *S*₆, one *S*₁₃ and one *S*₁₄ cDNA clones, and they predict closely related polypeptides. The *S*₆-derived sequence, referred to as BOS6 (Fig. 1), contains a continuous open-reading frame of 436 amino acids. In addition to including the full SLSG-coding regions, this sequence contains three residues omitted from the previously published sequence (see corrigendum, ref. 7).

A hydropathy profile generated from the predicted amino-acid sequence reveals a hydrophobic domain in the 31-residue N-terminal region of the polypeptide (Fig. 2). A molecule with the recognition function ascribed to SLSG is expected to be transported across membranes to the cell surface, so we suggest that this hydrophobic domain represents a signal peptide. BOS6 would therefore contain the full amino-acid sequence of SLSG, and the methionine residue at the beginning of the hydrophobic region (position –31) would represent the initiating codon for protein translation. This view is supported by the recently reported N-terminal amino-acid sequences of three SLSGs from *B. campestris*⁵ which are very similar to amino-acid residues 1–29 of our predicted sequence (Fig. 3a).

That no transmembrane domain is evident in the SLSG sequence (Fig. 2) is perhaps not surprising as SLSGs are water soluble. We have already shown that SLSGs are synthesized in the surface papillar cells of the stigma³, and because the interaction of self-incompatibility is known to take place at the pollen/papillar interface, we suggest that these molecules are transported to the papillar cell surface where they are deposited along the cell wall or in the periplasmic space between the cell membrane and the cell wall, rather than in the membrane itself.

The *S*-allele-specific differences in the electrophoretic mobilities of SLSGs following isoelectric focusing (IEF) and sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) are well documented^{2,3,8}, and are thought to reflect functional specificity-determining differences. In an attempt to deduce the possible structural basis of these differences, we compared the predicted amino-acid sequence of SLSGs encoded by the *S*₆, *S*₁₃ and *S*₁₄ alleles (Fig. 3b). The *S*₁₃ and *S*₁₄ cDNA clones are referred to as BOS13 and BOS14 respectively. The cDNA inserts of BOS13 and BOS14 end one and three amino-acid residues 3' to the presumptive initiating methionine codon respectively, but contain otherwise a full coding region. As expected for allelic variants, extensive similarity—approximately 80% at the protein level—is evident between the three sequences. They differ however by a number of amino-acid substitutions throughout their length, and by the deletion or addition of one or two amino acids at positions indicated by dashes in Fig. 3b.

Among other variations likely to be of potential importance in determining *S*-allele specificity are changes in potential sites for N-glycosylation and for serine- or threonine-linked O-glycosylation. The occurrence and the structure of N-linked carbohydrate chains in SLSG have recently been reported⁹. In all, eleven potential N-glycosylation sites which follow the general rule of Asn-X-Thr/Ser (ref. 10), where X is any residue, can be identified in the sequence (Fig. 2). Among these, only the sites at residues 15, 83, 90, 230, 284 and 359 are conserved in all three genotypes, whereas each *S*-allele specifies a distinct pattern of potential glycosylation sites at residues 33, 119, 214, 251 and 296.

A striking feature of the amino-acid sequences predicted from BOS6, BOS13 and BOS14, is the alternation in the SLSG polypeptide sequence of constant and variable regions (Figs 2 and 3b). The N-terminal region (A in Fig. 2), comprising amino-acid residues 1–181, is relatively conserved (144/181, or 79.6%, conserved positions). A second region (B in Fig. 2), residues 182–274, is the most variable with only 41.9% conservation (39/93 conserved positions). This is followed by a third conserved region (C in Fig. 2), residues 275–377 (83/103, or 80.6% conserved), which includes 10 invariant cysteine residues. Finally, the

§ Present address: 403 Althouse Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA.

32 62
 AG ATG AAA GGA GTA AGA AAA CCC TAC GAT AAT TCT TAT ACC TTA TCC TTT TTG CTT GTC TTT
 Met Lys Gly Val Arg Lys Pro Tyr Asp Asn Ser Tyr Thr Leu Ser Phe Leu Leu Val Phe
 -31
 92 122
 TTC GTC TTG ATY CTA TTT TGT CCT GCC TTT TCG ATC AAC ACT TTG TCG TCT ABA GAA TCT
 Phe Val Leu Ile Leu Phe Cys Pro Ala Phe Ser Ile Asn Thr Leu Ser Ser Thr Glu Ser
 1
 152 182
 CTT AGA ATC TCA AGC AAC AGA ACA CTT GTA TCT CCA GGT AAT AAC TTC GAA CTC GGC TTC
 Leu Arg Ile Ser Ser Asn Arg Thr Leu Val Ser Pro Gly Asn Asn Phe Glu Leu Gly Phe
 212 242
 TTC CGA ACC AAC TCA AGT TCT CGT TGG TAT CTC GGG ATA TGG TAC AAG AAA TTG CTC GAC
 Phe Arg Thr Asn Ser Ser Arg Trp Tyr Leu Gly Ile Trp Tyr Lys Lys Leu Leu Asp
 272 302
 AGA ACC TAT GTA TGG GTT GCC AAC AGA GAT AAC CCA CTC TCC AAT GCC ATT GGA ACC CTC
 Arg Thr Tyr Val Trp Val Ala Asn Arg Asp Asn Pro Leu Ser Asn Ala Ile Gly Thr Leu
 332 362
 AAA ATC TCA GGC AAT AAT CTT GTC CTC CTT GGT CAC ACC AAT AAA TCT GTT TGG TCG ACB
 Lys Ile Ser Gly Asn Asn Leu Val Leu Leu Gly His Thr Asn Lys Ser Val Trp Ser Thr
 392 422
 AAT CTT ACT AGA GGA AAT GAG AGA CTT CCG GTG GTG GCA GAG CTT CTC TCT AAT GGA AAC
 Asn Leu Thr Arg Gly Asn Glu Arg Leu Pro Val Val Ala Glu Leu Leu Ser Asn Gly Asn
 452 482
 TTC GTG ATG CGA GAC TCC AGT AAC AAC GAC GCA AGT GAA TAC TTG TGG CAA AGT TTC GAT
 Phe Val Met Arg Asp Ser Ser Asn Asn Asp Ala Ser Glu Tyr Leu Trp Gln Ser Phe Asp
 512 542
 TAC CCT ACG GAT ACT TTG CTT CCA GAG ATG AAA CTG GGT TAC GAC CTC AAA ACA GGG TTG
 Tyr Pro Thr Asp Thr Leu Leu Pro Glu Met Lys Leu Gly Tyr Asp Leu Lys Thr Gly Leu
 572 602
 AAC AGG TTC CTT ACA TCA TGG ABA AGT TCA GAT GAT CCA TCA AGC GGG GAT TTC TCG TAC
 Asn Arg Phe Leu Thr Ser Trp Arg Ser Ser Asp Asp Pro Ser Ser Gly Asp Phe Ser Tyr
 632 662
 AAG CTC GAA ACC CGA AGC CTT CTT GAG TTT TAT CTA TGG CAT GGG ATC TTT CCA ATG CAT
 Lys Leu Glu Thr Arg Ser Leu Pro Glu Phe Tyr Leu Trp His Gly Ile Phe Pro Met His
 692 722
 CCG AGT GGT CCA TGG AAT GGA GTC CCA TTT AGT GGC ATA CCA GAG GAC CAA AAG CTG AGT
 Arg Ser Gly Pro Trp Asn Gly Val Arg Phe Ser Gly Ile Pro Glu Tyr Leu Gln Lys Leu Ser
 752 782
 TAC ATG GTG TAC AAT TTC ACA GAG AAT AGT GAA GAG GTC GCT TAT ACA TTC CCA ATG ACC
 Tyr Met Val Tyr Asn Phe Thr Glu Asn Ser Glu Glu Val Ala Tyr Thr Phe Arg Met Thr
 812 842
 AAC AAC ACC ATC TAC TCG AGA TTG ACA CTA AGT TCC GAA GGG TAT TTT CAG CAG CTT ACG
 Asn Asn Ser Ile Tyr Ser Arg Leu Thr Leu Ser Ser Glu Gly Tyr Phe Gln Arg Leu Thr
 872 902
 TGG AAT CCG TCA ATA GGG ATA TGG AAC AAG TTC TGG TCT TCT CCA GTG GAC CCC CAG TGC
 Trp Asn Pro Ser Ile Gly Ile Trp Asn Arg Phe Trp Ser Ser Pro Val Asp Pro Gln Cys
 932 962
 GAT ACA TAC ATA ATG TGC GGG CTT TAC GCT TAC TGT GGC GTG AAC ACA TCA CTT GTT TGT
 Asp Thr Tyr Ile Met Cys Gly Pro Tyr Ala Tyr Cys Gly Val Asn Thr Ser Pro Val Cys
 992 1022
 AAC TGT ATC CAA GGG TTC AAT CCC CGA AAT ATA CAG CAG TGG GAT CAG AGA GTC TGG GCA
 Asn Cys Ile Gln Gly Phe Asn Pro Arg Asn Ile Gln Gln Trp Asp Gln Arg Val Trp Ala
 1052 1082
 GGT GGG TGT ATA AAG AGG ACG CCG CTT AGC TGC AGT GGA GAT GGT TTT ACA AGG ATG AAG
 Gly Gly Cys Ile Arg Arg Thr Arg Leu Ser Cys Ser Gly Asp Gly Phe Thr Arg Met Lys
 1112 1142
 AAC ATG AAG CTG CCA GAA ACT ACG ATG GCG ATT GTC GAC CCG AGT ATT GGT GTG AAA GAA
 Asn Met Lys Leu Pro Glu Thr Thr Met Ala Ile Val Asp Arg Ser Ile Gly Val Lys Glu
 1172 1202
 TGT GAG AAG AAG TGC CTT AGC GAT TGT AAT TGT ACT GCT TTT GCA AAT CCG GAT ATC CCG
 Cys Glu Lys Arg Cys Leu Ser Asp Cys Asn Cys Thr Ala Phe Ala Asn Ala Asp Ile Arg
 1232 1262
 AAT GGT GGG ACG GGT TGT GTG ATT TGG ACC GGA CCG CTT GAC GAT ATG CCG AAT TAC GTT
 Asn Gly Gly Thr Gly Cys Val Ile Trp Thr Gly Arg Leu Asp Asp Met Arg Asn Tyr Val
 1292
 GCT CAC GGT CAA GAT CTT TAT GTC AGA TTG GGT GTT GCT GAC CTT GTT TAG CTCCTTCCTT
 Ala His Gly Gln Asp Leu Tyr Val Arg Leu Ala Val Ala Asp Leu Val End
 AAAAAAAGCAGGATCC

Fig. 1 Nucleotide sequence and deduced amino-acid sequence of the cDNA insert in BOS6. The predicted amino-acid sequence is given below the nucleotide sequence and is numbered beginning with 1 for the first amino acid of the mature protein. The presumptive signal peptide sequence is indicated by negative numbers, and the potential polyadenylation signal is underlined.

Methods. The cDNA libraries were constructed from poly(A)⁺ RNA isolated from stigmas of a *B. oleracea* line homozygous for the *S*₆ allele, essentially as described⁴ except that the vector λ gt10 was used instead of pBR322. Recombinant phage were screened with the SLGS-encoding cDNA clone described previously⁴. The two clones with the largest homologous inserts were subcloned into pUC8, the restriction map of the inserts was determined, and overlapping fragments generated by digestion with two restriction enzymes were subcloned into the M13 vectors mp18 and mp19 for sequence determination in both orientations by the dideoxy chain-termination method¹¹ using a synthetic 17-mer as primer (New England Biolabs). For fragments generated by single restriction enzyme digestion, single-stranded DNA templates were subjected to 'C-tests' and different clones containing the two opposite strands were sequenced. For less accessible regions, *Sau* 3A shot-gun subcloning into M13 was used to complete the sequence.

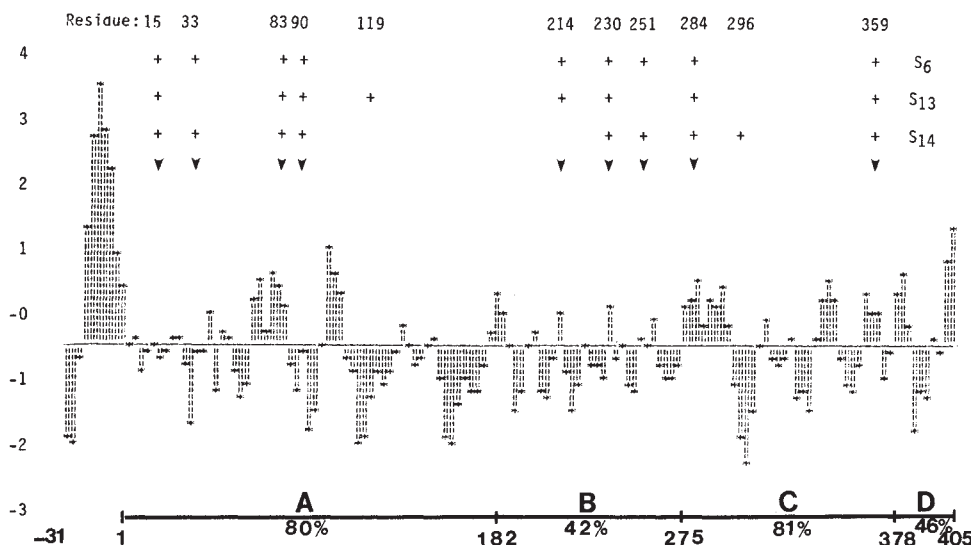


Fig. 2 Hydropathy profile of SLGS polypeptide predicted from the BOS6 sequence showing a hydrophobic signal peptide region at the N-terminus (residues -31 to 1). The profile was generated using the parameters of Kyte and Doolittle¹². Values above and below the horizontal line, hydrophobic and hydrophilic regions respectively; arrowheads, potential *N*-glycosylation sites in the BOS6 sequence; +, potential *N*-glycosylation sites along the sequences derived from the *S*₆, *S*₁₃ and *S*₁₄ alleles. A, B, C and D refer to regions showing approximately 80%, 42%, 81% and 46% conservation respectively. The highly variable region B (residues 182-274) is relatively hydrophilic. Region C (residues 275-378) contains 10 conserved cysteine residues.

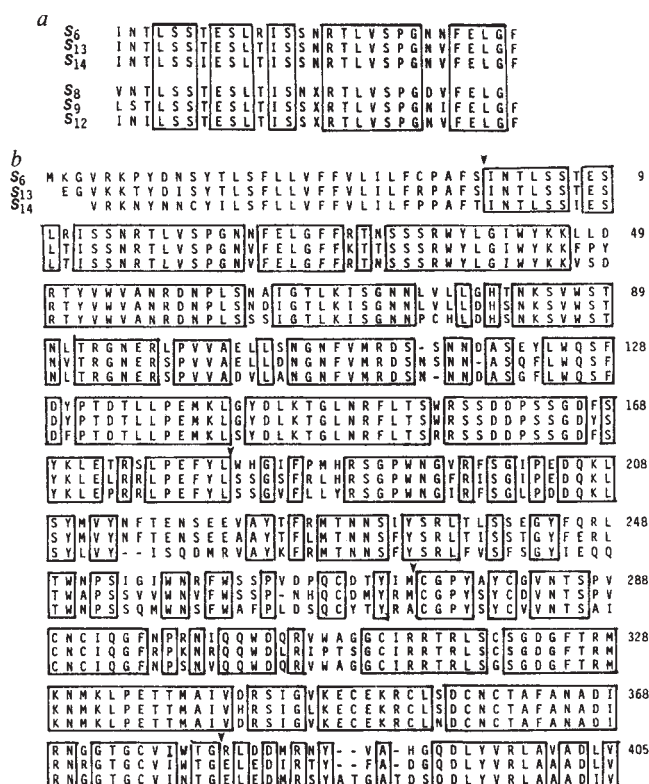


Fig. 3 Comparison of the amino-acid sequences of SLSGs from different *S*-alleles of *Brassica*. *a*, Homology of SLSGs at the N-terminus of the mature protein. The sequence of residues 1–29 of SLSGs encoded by the *S*₆, *S*₁₃ and *S*₁₄ alleles of *B. oleracea* were derived from the cDNA sequence analysis reported here, and the amino-acid sequences of SLSGs encoded by the *S*₈, *S*₉ and *S*₁₂ alleles of *B. campestris* were determined using a gas phase protein sequenator⁵. Boxes outline identical residues. *b*, Amino-acid sequences of SLSGs encoded by the *S*₆, *S*₁₃ and *S*₁₄ alleles of *B. oleracea*. (The nucleotide sequence of BOS13 and BOS14 are not presented but are available on request.) Identical residues are boxed, starting with residue 1 of the mature protein. Dashes, gaps introduced to maximize similarity. Arrowheads delineate the conserved and variable regions of the protein. The amino-acid sequence was predicted from the nucleotide sequences determined for both strands of the cDNA inserts of λ gt10 clones derived from the *S*₆, *S*₁₃ and *S*₁₄ homozygotes as in Fig. 1.

stretch of 28 amino acids at the C-terminus (D in Fig. 2) shows reduced homology of only 46.4% (13/28 conserved positions). This variation in the level of conservation over the protein is highly significant (heterogeneity $\chi^2 = 54.37$, 3 d.f., $P < 0.001$).

It is tempting to speculate on the functional significance to the cellular interaction of incompatibility of three features of SLSG deduced from the sequences presented. The *N*-glycosylation pattern typical of each allele and the highly variable region of the molecule could impart allelic specificity through the display of unique arrays of carbohydrate chains and/or amino-acid sequences. Note that the highly variable region is relatively hydrophilic (Fig. 2), and may therefore be exposed on the surface of the SLSG molecule. On the other hand, the conserved cysteine-rich region would provide for the intra- and inter-molecular bonds required to establish the tertiary, and possible quaternary configurations, necessary for the functioning of SLSGs in the recognition of pollen components.

This work was supported by a grant (no. 85-CRCR-1-1600) from USDA to J.B.N., and an NSF grant (no. DCB-8416301) to M.E.N. We thank Dr C. Aquadro for improving the manuscript. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00268.

Received 12 November 1986; accepted 12 February 1987.

1. Nishio, T. & Ninata, K. *Heredity* **38**, 391–396 (1977).
2. Nasrallah, J. B. & Nasrallah, M. E. *Experientia* **40**, 279–281 (1984).
3. Nasrallah, J. B., Doney, R. C. & Nasrallah, M. E. *Planta* **165**, 100–107 (1985).
4. Nasrallah, J. B., Kao, T.-H., Goldberg, M. L. & Nasrallah, M. E. *Nature* **318**, 263–267 (1985).
5. Takayama, S. *et al. Agric. Biol. Chem.* **50**, 1365–1367 (1986).
6. Anderson, M. A. *et al. Nature* **321**, 38–44 (1986).
7. Nasrallah, J. B., Kao, T.-H., Goldberg, M. L. & Nasrallah, M. E. *Nature* **326**, 523 (1987), (corrigendum).
8. Nishio, T. & Hinata, K. *Genetics* **100**, 641–647 (1982).
9. Takayama, S. *et al. Agric. Biol. Chem.* **50**, 1673–1676 (1986).
10. Neuberger, A., Gottschalk, A., Marshall, R. D. & Spiro, R. G. in *The Glycoproteins: Their Composition, Structure and Function Part A* (ed. Gottschalk, A.) 450–490 (Elsevier, Amsterdam, 1972).
11. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Rose, B. *J. molec. Biol.* **143**, 161–178 (1980).
12. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132.

Myristic acid is coupled to a structural protein of polyoma virus and SV40

Charles H. Streuli & Beverly E. Griffin

Department of Virology, RPMS, Hammersmith Hospital, London W12 0HS, UK

In the lytic cycle of papova viruses, both uncoating of the viral genome after infection and assembly of functional virions take place in the cell nucleus^{1,2}. The mechanisms by which newly internalized virions are targeted to the nucleus and viral DNA encapsidated into particles are poorly understood. Although the major capsid protein VP1 is involved in endocytosis³, and largely defines virion structure⁴, the functions of the minor proteins VP2 and VP3 have remained obscure. Here we show that VP2 from both polyoma virus and simian virus 40 (SV40) is covalently linked to myristic acid; this is the first report of a myristylated protein in the nucleus and of a fatty acid being important in the structure of a non-enveloped virus. We consider the implications of this unusual modification on encapsidation and suggest that VP2 may be a scaffolding protein for virion assembly.

While examining the relationship between gene function and the sub-cellular localization of viral antigens⁵ we investigated protein modifications that might be relevant to membrane association. Evidence that a polyoma-virus-induced function was modified by post-translational acylation with the fatty acid, myristic (*n*-tetradecanoic) acid, came from a comparison of the profiles of radiolabelled proteins from virus-infected and uninfected 3T6 cells. Extracts of cells pulse-labelled with [³⁵S]methionine during late stages of infection contained the capsid antigen VP1 (relative molecular mass 45,000; *M*_r 45K) as the major virally encoded polypeptide. Similar extracts from [³H]myristic-acid-labelled cells contained an infection-specific 35K polypeptide as a major component of the total myristylated cellular protein (Fig. 1a). The similarity of the size of this protein in wild-type and mutant virus-infected cells indicated that it was either a capsid antigen or a virally induced host protein. To distinguish between these possibilities, extracts were immuno-precipitated (Fig. 1b) with polyclonal antibodies raised against polyoma virions (anti-V serum). This serum, which specifically precipitates viral antigens from infected cells (lanes 1, 2, 3), recognized much of the available 35K protein from lysates of cells labelled with [³H]myristic acid (lanes 4, 5, 6). As VP2 of polyoma virus has an apparent relative molecular mass of 35K (ref. 6), it appeared a likely candidate for the myristylated species.

To test this, the protein composition of metabolically labelled polyoma virions was examined: [³⁵S]methionine-labelled particles consist of the viral capsid antigens VP1, VP2 and VP3, with VP2 appearing as a characteristic doublet^{6,7} (Fig. 2a); the major 35K myristylated species co-migrated with VP2 and was also visible as a doublet (Fig. 2b). A strong association between the

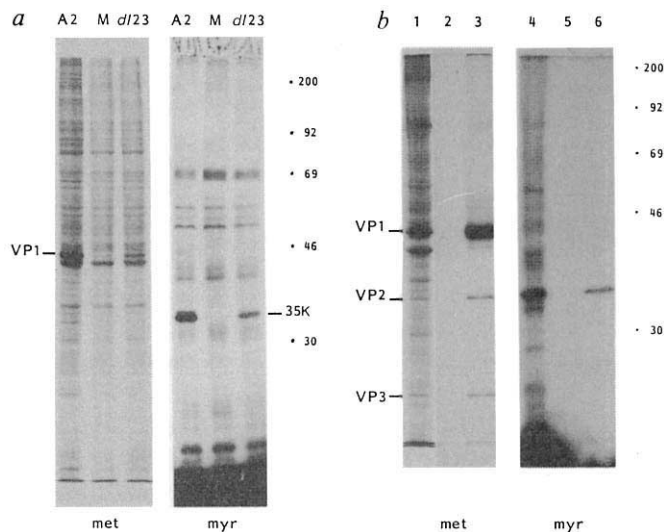


Fig. 1 Major viral proteins produced after infection of permissive cells. *a*, Swiss mouse 3T6 cells, infected (for 30 h) with wild-type (A2 strain) or *dl23* (a viral mutant with a 102 base-pair deletion in the early region which encodes truncated large-T and middle-T antigens³⁵) polyoma virus (20 plaque-forming units (p.f.u.) per cell), or mock-infected (M), were labelled (at 1 mCi per 10⁶ cells) for 4 h with [³⁵S]-methionine (met) or [³H]myristic acid (myr). Aliquots from total cell lysates prepared with a Nonidet P40 (NP40) lysis buffer (100 mM NaCl; 100 mM Tris HCl, pH 8.8; 0.5% NP40, ref. 36) and separated by SDS-polyacrylamide gel electrophoresis (PAGE). *b*, Wild-type virus-infected cells lysed and equal aliquots either separated directly by SDS-PAGE (lanes 1, 4), or immunoprecipitated in the presence (lanes 3, 6) or absence (lanes 2, 5) of anti-virus serum (M. A. Bioproducts) and *Staphylococcus aureus* before electrophoresis, as described³⁶. Lanes 2 and 3 were exposed for nine times as long as lane 1. The *M_r* of [¹⁴C]methylated marker proteins (in thousands) are shown in the margins (Amersham). Gels containing myristic-acid-labelled proteins were enhanced with Amplify (Amersham) before autoradiography.

[³H]myristic acid residue and VP2 was shown by the observation that the radiolabelled material in unfixed SDS gels (identical to that in Fig. 2b) could not be extracted with organic solvents (chloroform:methanol, 2:1 v/v; ref. 8). Neither hydroxylamine treatment (1 M in H₂O, pH 9.0; refs 9, 10) nor alkaline hydrolysis (0.2 M NaOH in absolute methanol⁸) released any isotope from VP2 in these gels (data not shown). Together, the data suggest that the fatty acid moiety is covalently linked to VP2 through an amide bond. To confirm the identity of the attached fatty acid, VP2 from purified polyoma and SV40 (see below) virions (metabolically labelled with [³H]myristic acid) was hydrolysed overnight with 6 M HCl and the resulting organic extract chromatographed on Whatman KC₁₈ thin-layer plates using acetonitrile:acetic acid (1:1, v/v) as a separating solvent¹¹. The released fatty acid co-migrated with a myristic acid marker (data not shown).

Comparison of the VP2 amino-acid sequences of six papova viruses reveals two main regions of homology (Fig. 3), one at the amino terminus and the other at the carboxy terminus. Amino-terminal methionine-glycine appears in five apparently unrelated myristylated proteins (the catalytic subunit of cyclic AMP-dependent protein kinase¹², calcineurin B¹³, mammalian retroviral gag proteins^{11,14}, NADH-cytochrome *b₅* reductase¹⁵ and p60^{v-src} (ref. 16)) suggesting that this sequence, which occurs in VP2 of all the papova viruses, may be a consensus sequence for myristyl modification. To explore this possibility, the protein profiles of SV40 virions were examined (Fig. 2c,d), and VP2, which in the case of SV40 migrates at 41K, was shown to be the only capsid antigen to be modified by myristic acid. It has

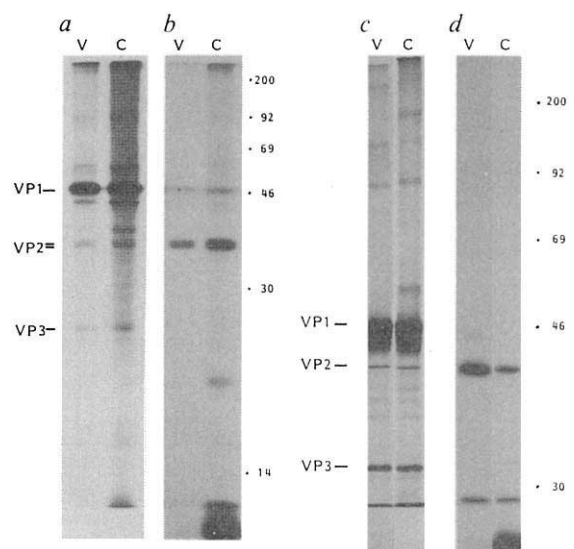


Fig. 2 Protein composition of polyoma and SV40 virus particles. *a*, *c*, [³⁵S]methionine or *b*, *d*, [³H]myristic acid-labelled proteins from purified polyoma (*a* and *b*), or SV40 (*c* and *d*), virions. Complete virion (V) and empty capsid (C) fractions are shown. **Methods.** Polyoma virus was purified from whole mouse embryo cells, infected with A2 virus (0.05 p.f.u. per cell) for 7 days, by established procedures³⁷. CV1 cells were used for SV40 production³⁷. Metabolic labelling was carried out in the presence of 1 mCi of appropriate label per dish (600 cm²) for the last 3 days of infection. Virions and capsids (buoyant density = 1.33 g cm⁻² and 1.28 g cm⁻², respectively) were separated by CsCl centrifugation³⁸, fractions collected, aliquots microdialysed, boiled in 100 mM Tris HCl, pH 6.8; 1% SDS; 1% mercaptoethanol; 0.01% bromophenol blue; 20% glycerol, and separated by SDS-PAGE; representative fractions corresponding to virus (V) and capsid (C) are as marked. The profile of [³⁵S]methionine-labelled proteins reflects the pattern observed after Coomassie blue staining (data not shown).

previously been found that the amino-terminal fragment of VP2 synthesized *in vivo* from polyoma virus could not be sequenced by Edman degradation¹⁷, and no methionine-containing tryptic fragments could be identified from this region¹⁸. In two of the myristylated proteins identified to date (p60^{v-src} and Pr65^{gag}), substitution of the glycine residue at the amino terminus has been shown to prevent modification^{19,20}. Taken together, the data suggest that the glycine at the amino terminus of VP2 of polyoma virus and SV40 is involved in its myristylation; the sequence comparison (Fig. 3) further suggests that the VP2s of other papova viruses might be similarly modified.

To investigate the subcellular location of the myristylated VP2, two procedures were used for preparing nuclei of infected cells^{21,22} (Fig. 4). Most of the myristylated VP2 was associated with the nuclear fraction after Dounce homogenization (compare lanes P1 and S1). Furthermore, mild detergent solubilized only about 5% of the VP2 from whole cells (lanes N and S) or from nuclei (lanes P1/N and P1/S). Analogous experiments using [³⁵S]methionine showed that the distribution of myristylated VP2 is representative of total VP2 (data not shown). Myristylated VP2 therefore appears to be present in cell nuclei—an expected result, given that isolated papova virions contain the modified protein—providing the first example of a nuclear myristylated protein (one previous study failed to detect significant acylated protein in the nuclear fraction of normal rat fibroblasts¹⁰). Most of the acylated proteins in total extracts of mammalian cells are membrane-associated^{10,23}. Of the characterized myristylated proteins that have been described (see refs 12–16, and ref. 24) all but two^{12,13} are membrane-bound, with

Fig. 3 Amino-acid sequence comparison of papova virus VP2s. About 40 residues from the amino and carboxy termini of polyoma VP2 (ref. 39) are compared with those of all the other published papova virus VP2 sequences (BKV³⁴, JCV⁴⁰, SV40³² hamster papova virus (HaPV)⁴¹ and lymphotropic papova virus (LPV)⁴²). Amino acids with homology to polyoma are shown in upper case. *, Positions where the amino-acid residue is identical throughout the species. O, Additional functionally homologous amino acids. Note the conserved glycine at the second residue. There is no further complete conservation of sequence for 15 residues and subsequent matches (not shown, for the sake of clarity) are sporadic until a region close to the carboxy terminus is reached. The amino-acid identities overall are restricted to 14%, but hydropathy plots (not included) show remarkable similarity of structure with the amino-terminal portions of each VP2 being entirely hydrophobic.

```

JCV  NH2--MGAALalLgDLvatvsEaaaaTGfSvaeIaaGEAAatiev----242aa----
BKV  NH2--MGAALalLgDLvasvsEaaaaTGfSvaeIaaGEAAatiev----241aa----
SV40 NH2--MGAALtLlLgDLiatvsEaaaaTGfSAEAILSGEAAAAiev----242aa----
PY    NH2--MGAALTILVDLIEGLAEVSTLTGLSAEAILSGEALALDg----237aa----
HaPV  NH2--MGAaIaivliemIEGLAEVSTLTGLSAEAILSGEAAfAAIda----262aa----
LPV   NH2--MGaVLsILfnIaEiaAEISlEtGTftvDAiltGEAAfAAvet----244aa----
      **  oo oo oo  o  o  **ooo  *o  ****o  o

```

```

QRsaPqWMLPLlLGLYgtvTPa----leay---EDGPnKKRRl----21aa---COOH
QRtaPqWMLPLlLGLYgtvTPa----leay---EDGPnKKRRl----28aa---COOH
QRtaPqWMLPLlLGLYgsvTsa----lkay---EDGPnKKRRl----28aa---COOH
QRVTPDWMLPLlLGLYGDITPTWATVid---ev-EyGpKKRRf----1aa---COOH
QRVTPDWMLPLlLGLYGDITPTWATVid---ev-EyGpKKRRf----1aa---COOH
QRVTPDWMLPLlLGLYGDITPTWATVid---ev-EyGpKKRRf----1aa---COOH
QRVTPDWMLPLlLGLYGDITPTWATVid---ev-EyGpKKRRf----1aa---COOH
      **  ****o****o****o  o  *  *  o****o

```

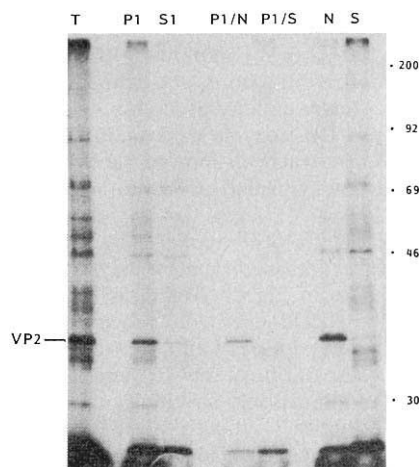


Fig. 4 Subcellular distribution of myristylated VP2. Cells were infected, labelled with [³H]myristic acid (as in Fig. 1 legend) and aliquots from different fractions separated by SDS-PAGE. T, total cell extract, sonicated in RIPA buffer (150 mM NaCl; 10 mM Tris/HCl, pH 8.8; 1% deoxycholic acid (DOC); 1% NP40; 0.1% SDS, ref. 43). P1 and S1, cells swollen in hypotonic buffer (20 mM HEPES, pH 7.1; 5 mM KCl; 3 mM MgCl₂), Dounce homogenized and the nuclei pelleted and washed once at 1,000g (ref. 21). P1 is the nuclear pellet (also containing large membrane sheets) and S1 the soluble component and all non-sedimentable vesicles. N and S, cells lysed in NP40 'cytoskeleton' buffer (10 mM PIPES, pH 6.8; 100 mM NaCl; 3 mM MgCl₂; 300 mM sucrose; 0.5% NP40), the insoluble nuclear fraction, including cytoskeleton (N) and detergent-soluble material containing most membrane components (S) separated by centrifugation at 200g (ref. 22). P1/N and P1/S, Dounce prepared nuclei, detergent-washed (as for samples N and S) and separated into the residual nuclei (P1/N) and soluble (P1/S) fractions. In each case, the VP2 from the total extract (shown) could be specifically immunoprecipitated with anti-V serum and *S. aureus* (not shown).

the fatty acid residue being essential for this association in the case of p60^{v-src} (ref. 25) and mammalian retrovirus Pr65^{gag} (ref. 20). Although the hydrophobicity of the amino terminus of VP2 (Fig. 3) and its association with a myristyl residue suggest a membrane location, the majority was insoluble in non-ionic detergents (Fig. 2 lanes N and S). As this indicates attachment to cytoskeletal structures (previously noted in the case of SV40^{26,27}) the myristyl-acyl chain of VP2 could influence its membrane orientation to form a stable association with the nucleoskeleton.

Our observation that VP2 is myristylated raises a number of questions about the role of this modification during the normal

lytic cycle of papova viruses: mutants with deletions in the region coding uniquely for VP2 have severely impaired growth²⁸, and recent data indicate that VP2 *per se* is essential for efficient production of both polyoma virus (R. Sahli and T. L. Benjamin, personal communication) and SV40 (L. A. Barber and J. D. Gralla, personal communication). Under conditions favouring crystallization *in vitro*, VP1 contains sufficient structural information for aggregation into pentameric capsids and subsequent self-assembly into empty viral capsomers²⁹. VP2 and VP3 may however, be necessary to catalyse rapid formation of true viruses *in vivo*. Both the involvement of these antigens (perhaps as scaffolding proteins) and post-translational modifications of VP1 (ref. 30), are probably required for efficient encapsidation of DNA into intact virions. Further, the carboxy terminus of VP2 contains a highly conserved hydrophobic sequence followed by a strongly hydrophilic domain (Fig. 3); the latter has a putative nuclear location signal³¹, and may be a region which can interact with viral DNA³²⁻³⁴. When correctly positioned in the nucleoskeleton, VP2 could anchor the viral minichromosome for nucleation of capsid assembly. In support of this notion, myristylation of Pr65^{gag} has been found to be essential for assembly of certain mammalian retroviruses²⁰. Studying variants of polyoma with specific mutations at the first glycine residue of VP2 (currently in progress) should confirm the site of modification and indicate whether this is necessary for the viral life cycle.

We thank Drs Martin Allday, Mick Jones, David King and John Sinclair for constructive criticisms, and the MRC for research funding.

Received 12 December 1986; accepted 19 February 1987.

- Mackay, R. L. & Consigli, R. A. *J. Virol.* **19**, 620-636 (1976).
- Toozee, J. *DNA Tumor Viruses* (Cold-Spring Harbor Laboratory, New York, 1981).
- Consigli, R. A., Griffith, G. R., Marriot, S. J. & Ludlow, J. W. in *Virus Attachment and Entry Into Cells* (eds Crowell, R. L. & Lonberg-Holm, K.) 44-53 (American Society for Microbiology Washington, DC, 1986).
- Rayment, I., Baker, T. S., Caspar, D. L. D. & Murakami, W. T. *Nature* **295**, 110-115 (1982).
- Dilworth, S. M. *et al. EMBO J.* **5**, 491-499 (1986).
- Hewick, R. M., Fried, M. & Waterfield, M. D. *Virology* **66**, 408-419 (1975).
- Gibson, W. *Virology* **62**, 319-336 (1974).
- Staufenbiel, M. & Lazarides, E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 318-322 (1986).
- Magee, A. I. *et al. Biochim. biophys. Acta* **798**, 156-166 (1984).
- Magee, A. I. & Courtneidge, S. A. *EMBO J.* **4**, 1137-1144 (1985).
- Schultz, A. & Oroszlan, S. *Virology* **133**, 431-437 (1984).
- Carr, S. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6128-6131 (1982).
- Aitken, A. *et al. FEBS Lett.* **150**, 314-318 (1982).
- Henderson, L. E., Krutzsch, H. C. & Oroszlan, S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 339-343 (1983).
- Ozols, J., Carr, S. A. & Strittmatter, P. *J. biol. Chem.* **259**, 13349-13354 (1984).
- Schultz, A. M. *et al. Science* **227**, 427-429 (1985).
- Hewick, R. M., Mellor, A., Smith, A. E. & Waterfield, M. D. *J. Virol.* **33**, 631-636 (1980).
- Hunter, T. & Gibson, W. *J. Virol.* **28**, 240-253 (1978).
- Kamps, M. P., Buss, J. E. & Sefton, B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4625-4628 (1985).
- Rein, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 7246-7250 (1986).
- Manger, R., Rasheed, S. & Rohrschneider, L. *J. Virol.* **59**, 66-72 (1986).
- Capco, D., Wan, K. M. & Penman, S. *Cell* **29**, 847-858 (1982).
- Olson, E. N. & Spizz, G. *J. biol. Chem.* **261**, 2458-2466 (1986).

24. Marchildon, G. A., Casnellie, J. E., Walsh, K. A. & Krebs, E. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7679-7682 (1984).
25. Buss, J. E., Kamps, M. P., Gould, K. & Sefton, B. M. *J. Virol.* **58**, 468-474 (1986).
26. Lin, W., Hata, T. & Kasamatsu, H. *J. Virol.* **50**, 363-371 (1984).
27. Streckert, H.-J. *et al. J. cell. Biochem.* **31**, 277-287 (1986).
28. Cole, C. N. *et al. J. Virol.* **24**, 277-294 (1977).
29. Salunke, D. M., Caspar, D. L. D. & Garcea, R. L. *Cell* **46**, 895-904 (1986).
30. Garcea, R. L., Ballmer-Hofer, K. & Benjamin, T. L. *J. Virol.* **54**, 311-316 (1985).
31. Smith, A. E. *et al. Proc. R. Soc. B* **226**, 43-58 (1985).
32. Fiers, W. *et al. Nature* **273**, 113-120 (1978).
33. Llopis, R. & Stark, G. R. *J. Virol.* **38**, 91-103 (1981).
34. Seif, I., Khoury, G. & Dhar, R. *Cell* **18**, 963-977 (1979).
35. Smolar, N. & Griffin, B. E. *J. Virol.* **38**, 958-967 (1981).
36. Dilworth, S. M. & Griffin, B. E. *Proc. natn. Acad. Sci.* **79**, 1059-1063 (1982).
37. Turtler, H. & Beard, P. in *Virology, a Practical Approach* (ed. Mahy, B.) 169-192 (IRL Press, Oxford, 1985).
38. McMillen, J., Center, M. S. & Consigli, R. A. *J. Virol.* **17**, 127-131 (1976).
39. Arrand, J. R. *et al. J. Virol.* **33**, 606-618 (1980).
40. Frisque, R. J., Bream, G. L. & Cannella, M. T. *J. Virol.* **51**, 458-469 (1984).
41. Delmas, V., Bastien, C., Scherneck, S. & Feunteun *EMBO J.* **4**, 1279-1286 (1985).
42. Pawlita, M., Clad, A. & zurHausen, H. *Virology* **143**, 196-211 (1985).
43. Brugge, J. S. & Erikson, R. L. *Nature* **269**, 346-348 (1977).

The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses

Graeme J. Wistow*, John W. M. Mulders† & Wilfried W. de Jong†

* Laboratory of Molecular and Developmental Biology, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Department of Biochemistry, Centre of Eye Research, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, Netherlands

The major components of mammalian lenses are tissue-specific, soluble proteins, the α -, β - and γ -crystallins¹⁻³. The lenses of other vertebrate classes often contain other major proteins, notably δ -crystallin in birds and reptiles⁴. A fourth distinct type, described as ϵ -crystallin, is prominent in many bird and crocodile lenses^{5,6}. Here we show that ϵ -crystallin is an active glycolytic enzyme, lactate dehydrogenase (LDH) (EC 1.1.1.27) and that duck ϵ -crystallin appears to be identical to duck LDH-B4. LDH is a normal metabolic component in other lenses⁷, but in duck is present in amounts far exceeding the requirements of any likely catalytic role. It appears that an active enzyme has been recruited, unchanged, to an extra role as a structural protein in the lens without gene duplication and sequence divergence. This surprising discovery raises the possibility that other crystallins may similarly be enzymes expressed at high levels in lens as structural proteins.

Lens proteins occupy one of the more unusual environments in vertebrate biology. They have extraordinary requirements for structural stability, because they must be exposed to light and yet survive at high concentration (crystallins account for 20-50% of the wet weight of lenses) without significant turnover for as long as the whole life of the organism^{1,3,8}. Because of the relatively recent evolutionary origin of the lens, it has been recognized that its protein components are probably derived from non-lens ancestors⁸. This has been verified for the α -crystallins which belong to the same protein superfamily as the small heat-shock proteins^{9,10}, whereas the β - and γ -crystallins are related to each other¹¹⁻¹³ and also to protein S, a calcium-binding protein of the bacterium *Myxococcus xanthus*¹⁴.

The α - and β -crystallins are represented in all vertebrate lenses: δ -crystallin⁴ is prominent in all birds and reptiles, ρ -crystallin is a major component of frog lenses¹⁵, and τ -crystallin is found in lampreys and in some birds, reptiles and fish^{16,17}. The lenses of many birds and crocodiles contain in addition to α -, β - and δ -crystallins another major protein, ϵ -crystallin, which accounts for up to 23% of the total lens protein (Fig. 1).

Partial peptide sequences for duck ϵ -crystallin (duck- ϵ)⁵ show close similarity with vertebrate lactate dehydrogenases (Fig. 2).

Over the ranges compared, (52% of the LDH sequence) duck- ϵ exhibits 94% identity with chicken LDH-B¹⁸ even though the comparison includes the highly variable amino-terminal region. The identity with chicken LDH-A¹⁸ is 70% while both duck- ϵ and chicken LDH-B show 85-88% identity with pig LDH-B¹⁹. There are a few noteworthy differences—Ser 20 replaces Asn, which is conserved in the other sequences, and Asn 114 and Phe 118 are replaced by glycines in duck- ϵ —but all the residues required for LDH activity in the sequences compared^{20,21} are present in duck- ϵ . Duck heart LDH-B4 isolated by standard methods^{22,23} shows immunocrossreactivity with duck- ϵ (Fig. 1) and other physical parameters for duck- ϵ ⁵ are also consistent with a close relationship to LDH-B4, a normal glycolytic enzyme of other tissues.

Structural similarity alone does not necessarily imply enzymatic function. For example, haptoglobin is a protein evolutionarily related to serine proteinases but lacking the ancestral enzyme activity²⁴. However, duck- ϵ does have LDH activity. Duck lens extract, assayed by standard methods^{22,23}, has 180 times the activity of chicken lens extract, and isolated duck- ϵ has an activity of 100 U mg⁻¹ compared with 150 U mg⁻¹ for duck LDH-B4 and both exhibit inhibition at high concentrations of pyruvate^{22,23}. The lower activity of duck- ϵ may be the result of ageing in the lens⁷. On isozyme electrophoresis²⁵ duck and chicken lens and heart extracts all showed the same single band of LDH activity with the typical avian low anodal mobility (not shown).

Duck LDH-B and duck- ϵ were compared by peptide mapping (Fig. 3). The only differences were changes in charge for peptides 5 and 34 (numbered as in Fig. 2). Amino-acid analysis of the LDH-B peptides showed identity with the equivalent duck- ϵ peptides. The changes in peptides 5 and 34 probably result from deamidation (Asn $\rightarrow$ Asp) at positions 163 and 265, a common post-translational modification. Strikingly, all the unusual sequence differences between chicken LDH-B and duck- ϵ are also observed in duck LDH-B including changes such as Asn 114 $\rightarrow$ Gly and Leu 332 $\rightarrow$ Met which are not seen in heron and alligator ϵ -crystallin⁵.

It could be that ϵ -crystallin is the result of a recent duplication of the LDH-B gene. However, it must have been present in the common ancestor of birds and crocodiles at least 240 Myr ago²⁶ and would be expected to show greater divergence from LDH-B than is apparent. That duck- ϵ has LDH activity also argues against a gene-duplication hypothesis. The extremely high LDH activity in duck lens greatly exceeds any plausible glycolytic requirements, implying that duck- ϵ is not required in lens purely for its enzymatic activity. If it has a non-enzymatic role there should be no selective pressure to retain the ancestral enzyme activity which, as in the case of haptoglobin²⁴, would be lost.

Overall these results suggest that ϵ -crystallin is the product of the same gene as LDH-B4 which in some birds and crocodiles is subject to two different patterns of tissue-specific expression. First, as in other species, it is differentially expressed between such tissues as heart and skeletal muscle as a metabolic enzyme²³. Secondly, it is expressed at very high levels in the lens.

It is tempting to speculate that LDH has been recruited for some functional characteristic. Apart from its role in glycolysis, LDH can be seen as an NAD⁺/NADH binding protein, with high affinity for NADH²³, which has an absorption maximum at 340 nm. Bird retinas are sensitive at 370 nm in the ultra-violet²⁷ and NADH bound in the lens could usefully serve to reduce glare in the same way that other chromophores act at longer wavelengths in mammalian lenses³⁸. Indeed, the birds which possess detectable ϵ -crystallin (Fig. 1) all hunt or feed through or near water surfaces or at high altitude and could benefit from glare reduction. Alternatively, ϵ -crystallin/LDH could sequester NADH in the lens as an energy source or even as a reservoir of reducing potential: in the latter case filling a role which has been proposed for the sulphhydryl-rich γ -crystallins²⁹, which are absent in birds³⁰. However, enzymatic or binding functions may

Fig. 1 The occurrence of ϵ -crystallin in different species and comparison with LDH. Top, Coomassie blue-staining pattern following SDS gel electrophoresis of lens extracts^{1,5}. Lanes: 1, emu (*Dromaius novaehollandiae*, Casuariformes); 2, gannet (*Sula bassana*, Pelecaniformes); 3, stork (*Ciconia ciconia*, Ciconiiformes); 4, flamingo (*Phoenicopterus ruber*, Phoenicopteriformes); 5, buzzard (*Buteo buteo*, Falconiformes); 6, owl (*Asio otus*, Strigiformes); 7, budgerigar (*Melopsittacus undulatus*, Psittaciformes); 8, swift (*Apus apus*, Apodiformes); 9, pigeon (*Columba palumbus*, Columbiformes); 10, gull (*Larus ridibundus*, Charadriiformes); 11, sparrow (*Passer domesticus*, Passeriformes); 12, chicken (*Gallus domesticus*, Galliformes); 13, duck (*Anas platyrhynchos*, Anseriformes); 14, calf; 15, guinea pig; 16, rabbit; 17, frog (*Rana temporaria*); 18, alligator (*Alligator mississippiensis*); 19, duck ϵ -crystallin; 20, duck-heart LDH-B. The major crystallin polypeptides are indicated. The water-soluble fraction of duck lens contains 97% of the total lens protein. Bottom, identification of ϵ -crystallin on a corresponding immunoblot obtained by transfer to nitrocellulose, reaction with rabbit antiserum against duck ϵ -crystallin and peroxidase staining⁵. Each lane contained 10–20 μ g of lens protein. The amounts of ϵ -crystallin measured by densitometric scanning were: gannet, 23%; stork, 18%; flamingo, 12%; buzzard, 2%; swift, 5%; gull, 3%; duck, 10% and alligator, 8%. The bands in rabbit, guinea pig (Q.-L. Huang, P. Russell, S. H. Stone and J. S. Zigler, Jr, manuscript in preparation) and frog¹⁵ of relative molecular mass (M_r) 38,000 give no detectable immunoreaction with the ϵ -crystallin antiserum. Previously ϵ -crystallin has been demonstrated in all ducks and geese investigated and in loons (Gaviiformes), curlew (Ciconiiformes), oystercatcher (Charadriiformes) and was not detectable in woodpecker (Piciformes), cuckoo (Cuculiformes), coot (Gruiformes), penguin (Sphenisciformes) and fulmar (Procellariiformes)⁵. The presence of ϵ -crystallin has thus been examined in representatives of 19 of the 25 or so avian orders. The distribution does not appear to correspond to phylogenetic relationships between the avian orders³². In addition, ϵ -crystallin has been found in caiman, but not lizard, snake or turtle⁵.

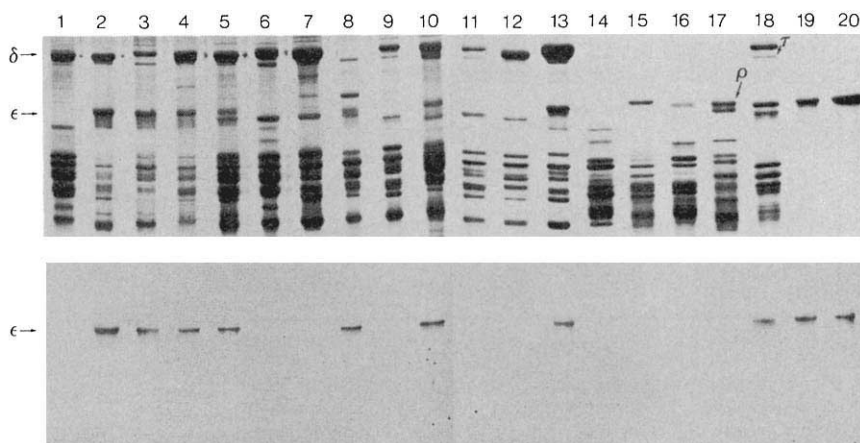
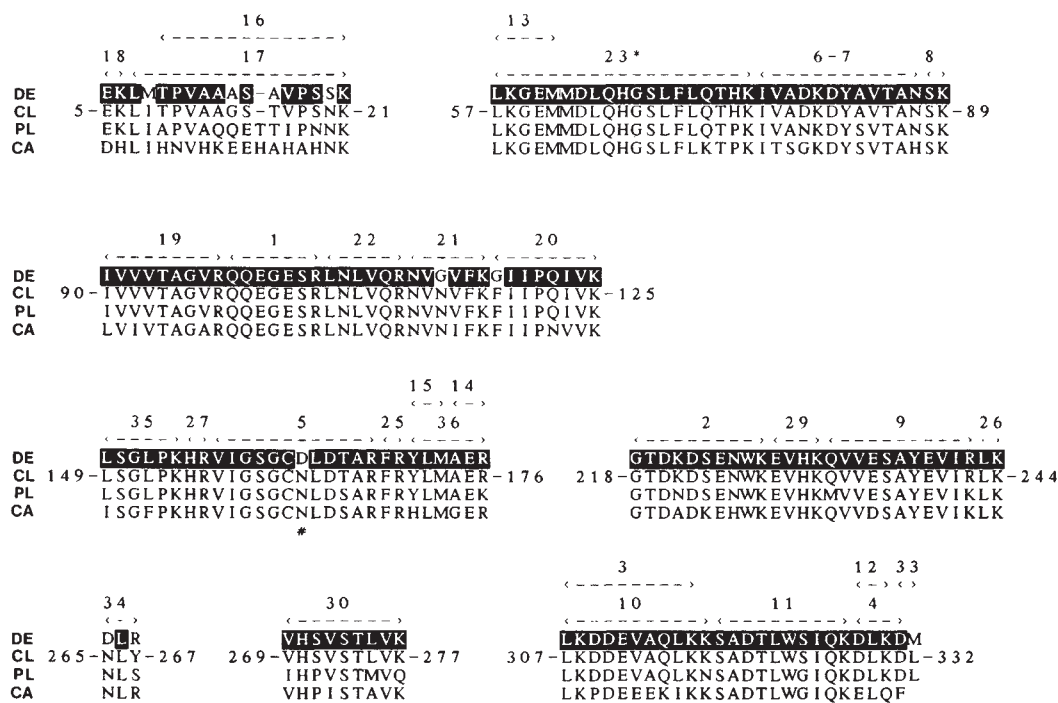


Fig. 2 Alignment of the amino-acid sequences of tryptic peptides of duck- ϵ (DE) (inferred from amino-acid composition and partial sequence determination⁵) with the equivalent regions of chicken (CL) and pig (PL) LDH-B and chicken LDH-A (CA)^{18,19}. Cleavage after Met is a result of prior CNBr treatment⁵. Single letter code is used. Sequence numbers are for the chicken LDH-B sequence. Numbers above the sequence and broken lines enclosed by arrows correspond to the peptides of duck- ϵ described earlier⁵ and to the peptides of duck- ϵ and duck LDH-B as isolated in this study (see Fig. 3). The duck- ϵ sequence is boxed where it matches chicken LDH-B. Residues 163 and 265 (positions marked by #) are apparently aspartates in duck- ϵ compared with asparagines in duck and chicken LDH-B, presumably the result of deamidation. The 174 residues compared represent 52% of the LDH sequences.



not be the reasons for the recruitment of LDH as a lens protein. LDH-B4 is known to be very stable²³ and its tightly associated tetrameric structure²¹ perhaps makes it suitable as a lens protein simply in terms of thermodynamic stability.

If one enzyme has been used in this way, other taxon-specific crystallins, such as δ - (which shows non-lens expression³¹), τ - and ρ -crystallins, may have similar origins.

Although α - and β -crystallins seem to be classic examples of proteins evolved from ancestors of different function and tissue-specificity by gene duplication and sequence divergence, ϵ -crystallin/LDH may illustrate a different evolutionary approach in which a functional gene acquires an additional role without duplication. At present the product of the putative ϵ -crystallin/Ldh-b gene is presumably subject to selective

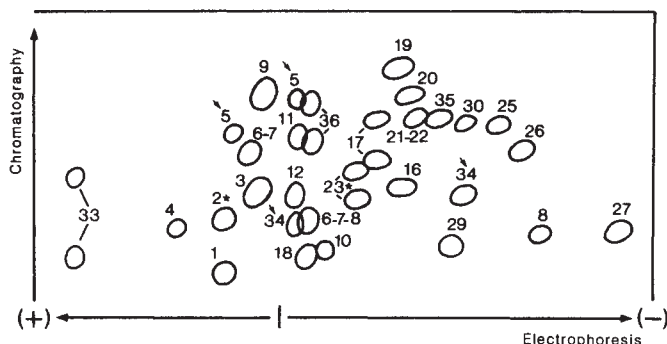


Fig. 3 Peptide maps of duck ϵ -crystallin and duck-heart LDH. Procedures were as described⁵, except that the CNBr treatment before tryptic digestion was omitted. Proteins (10 mg ml⁻¹) were digested with 1% (w/w) trypsin for 2 h at 37 °C in 0.1 M NH₄HCO₃, pH 8.6. The digestion was stopped by heating for 2 min at 100 °C and the insoluble peptides were removed by centrifugation. The soluble peptides were separated by high-voltage paper electrophoresis (pH 6.5) followed by descending paper chromatography. Peptides were localized with fluorescamine and eluted for amino-acid analysis. The numbers of the tryptic peptides correspond to those in Fig. 2, apart from peptide 2* (which is residues 222–227). Peptides 17, 23*, 33 and 36 contain methionine and were found in the unoxidized and oxidized forms with different R_f . The peptide maps and all amino-acid compositions were identical for ϵ -crystallin and duck-heart LDH apart from peptides 5 and 34 (arrows) which are more acidic in ϵ -crystallin. Unusual incomplete tryptic cleavage occurred after Met 8 (peptide 16) and Asn 87 (peptide 6–7).

pressure both as an enzyme and as a lens protein, perhaps explaining some of the unusual sequence changes in duck- ϵ compared with other vertebrate LDH subunits.

We thank Willeke Hoekman and Marlies Versteeg for technical assistance, Dr Steven Stapel for antisera and Dr Jan Wattel for lenses. This work was supported in part by the Netherlands Foundation for Chemical Research (SON) and the Netherlands Organisation for the Advancement of Pure Research (ZWO). We also thank Professor Hans Bloemendal and Dr Joram Piatigorsky for helpful discussions and support.

Received 5 December 1986; accepted 13 February 1987.

- Bloemendal, H. (ed.) *Molecular and Cellular Biology of the Eye Lens* (Wiley, New York, 1981).
- Piatigorsky, J. *Cell* **38**, 620–621 (1984).
- De Jong, W. W. & Hendriks, W. J. *molec. Evol.* **24**, 121–129 (1986).
- Piatigorsky, J. *Molec. cell. Biochem.* **59**, 33–56 (1984).
- Stapel, S. O., Zweers, A., Dodemont, H. J., Kan, J. H. & de Jong, W. W. *Eur. J. Biochem.* **147**, 129–136 (1985).
- Brahma, S. K. & Defize, L. H. *Curr. Eye Res.* **6**, 679–684 (1985).
- Hockwin, O. & Ohloff, C. in *Molecular and Cellular Biology of the Eye Lens* (ed. Bloemendal, H.) 367–413 (Wiley, New York, 1981).
- De Jong, W. W. in *Molecular and Cellular Biology of the Eye Lens* (ed. Bloemendal, H.) 221–278 (Wiley, New York, 1981).
- Ingolia, T. D. & Craig, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2360–2364 (1982).
- Wistow, G. *FEBS Lett.* **181**, 1–6 (1985).
- Driessen, H. P. C., Herbrink, P., Bloemendal, H. & de Jong, W. W. *Eur. J. Biochem.* **121**, 83–91 (1981).
- Blundell, T. *et al. Nature* **289**, 771–777 (1981).
- Wistow, G. *et al. FEBS Lett.* **133**, 9–16 (1981).
- Wistow, G., Summers, L. & Blundell, T. *Nature* **315**, 771–773 (1985).
- Gause, G. G., Tomarev, S. I., Zinovieva, R. D., Arutyunyan, K. G. & Dolgilevich, S. M. in *The Lens: Transparency and Cataract* (ed. Duncan, G.) 171–179 (EURAGE, Rijswijk, Netherlands, 1986).
- Stapel, S. O. & de Jong, W. W. *FEBS Lett.* **162**, 305–309 (1983).
- Williams, L. A., Ding, L., Horwitz, J. & Piatigorsky, J. *Expl Eye Res.* **40**, 741–749 (1985).
- Torff, H.-J., Becker, D. & Schwarzwaldner, J. in *Pyridine Nucleotide Dependent Dehydrogenases* (ed. Sund, H.) 31–42 (De Gruyter, Berlin, 1977).
- Kiltz, H.-H., Keil, W., Griesbach, M., Petry, K. & Meyer, H. *Hoppe-Seyler's Z. physiol. Chem.* **358**, 123–127 (1977).
- Eventoff, W. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 2677–2681 (1977).
- Grau, U. M., Trommer, W. E. & Rossmann, M. G. *J. molec. Biol.* **151**, 284–307 (1981).
- Stolzenbach, F. *Meth. Enzym.* **9**, 278–288 (1966).
- Holbrook, J. J., Liljas, A., Steindal, S. J. & Rossmann, M. G. in *The Enzymes* Vol. XI (ed. Boyer, P. D.) 191–292 (Academic, New York, 1975).
- Kurosky, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3388–3392 (1980).
- Rosalki, S. B. *Clin. Biochem.* **7**, 29–37 (1974).
- Benton, M. J. *Nature* **312**, 599 (1984).

- Chen, D.-M., Collins, J. S. & Goldsmith, T. H. *Science* **225**, 337–340 (1984).
- Walls, G. L. *The Vertebrate Eye and its Adaptive Radiation* (Hafner, London, 1963).
- Wistow, G. *et al. J. molec. Biol.* **170**, 175–202 (1983).
- Treton, J. A., Jones, R. E., King, C. R. & Piatigorsky, J. *Expl Eye Res.* **39**, 513–522 (1984).
- Clayton, R. M., Jeanny, J.-C., Bower, D. J. & Errington, L. H. in *Current Topics in Developmental Biology* Vol. 20 (eds Okada, T. S. & Kondoh, H.) 137–152 (Academic, New York, 1986).
- Cracraft, J. *Auk* **98**, 681–714 (1981).

Structure and evolution of ricin B chain

Earl Rutenber, Michael Ready & Jon D. Robertus

Clayton Foundation Biochemical Institute and Department of Chemistry, University of Texas, Austin, Texas 78712, USA

Ricin is a dimeric toxin from the castor bean *Ricinus communis*, which is composed of a sugar-binding subunit (B) that attaches to receptors on the surfaces of target cells^{1–4} and a subunit (A) with enzymatic activity that attacks and inactivates ribosomes. We report here that comparison of amino-acid sequence data with high-resolution structure analysis of the ricin B subunit shows it to be the product of a series of gene duplications. The modern protein has two sugar-binding domains, each of which is composed of three copies of a more ancient galactose-binding peptide of about 40 residues.

The B chain of ricin is known to bind two galactosides in a non-cooperative manner^{5,6}. Recent chemical work has implicated Tyr 248 in the 'strong' galactose binding site⁷, and an undetermined tryptophan in the 'weak' site⁸. We observed several years ago that the amino-acid sequence of the B chain could be broken into two homologous domains, which were presumed each to bind one galactoside⁹. That each domain contained two disulphide loops shown to be homologous to residues 168–225 of discoidin I, a galactose-binding protein from the slime mould *Dictyostelium discoideum*, suggested that the peptide was an ancient galactoside-binding unit¹⁰. We have recently solved the three-dimensional structure of ricin to 2.8 Å resolution by X-ray crystallography¹⁵. The structure shows that the B chain does indeed fold into two separate globular domains with identical folding topologies, each domain binding a lactose moiety. However, the lactose sugars were bound at apparently non-homologous sites in the two domains.

An analysis of the primary and tertiary structure of the B chain reveals that each of the homologous domains can be divided into four peptides, here called the λ -, α -, β - and γ -subdomains. The λ -peptide of domain 1, residues 1–16, has been previously shown to be homologous to the λ -peptide of domain 2 (peptide 2 λ), residues 136–150, but these are not related to the others. However, the α -, β - and γ -units which comprise the main body of each domain are all homologous with each other (Fig. 1a). The units have undergone considerable divergence, but the similarities are statistically significant (Fig. 1b). Key residues in the units are the invariant Trp and strongly conserved Ile residues marked with asterisks. The Asp and Gln residues of the galactose-binding site are marked by arrows.

The structure and interactions of the units in each domain are shown in Fig. 2, a schematic ribbon drawing of the B chain viewed down the crystallographic a axis¹⁵. The six units are highlighted and form the bulk of each domain. One of the linking peptides joins the two domains (2 λ) and the other connects the B chain to the A chain (1 λ).

The structural similarity of the α -, β - and γ -subdomains to each other is seen in Fig. 3, where they are placed in common orientation. The positions of bound lactose molecules in subdomains 1 α and 2 γ are also shown, demonstrating that the two binding sites, which are not homologous at the domain level, are indeed homologous at the level of the subdomains. Each

Fig. 1 *a*, Amino-acid sequence alignment of the subdomain units of ricin B chain. The six sequences from ricin are aligned from amino-toward carboxy-terminus and are arranged to maximize their similarities. The subdomain units correspond to residues as follows: 1 α residues 17–59; 1 β , 60–100; 1 γ , 103–135; 2 α , 151–183; 2 β , 187–226; 2 γ , 228–262. The bottom line shows residues 124–143 of *E. coli* galactose-binding protein (GBP). Boxes enclose identical residues. Sequences are as determined by Halling *et al.*¹⁴. *b*, Statistical relationships among the B chain subdomains. The relationship among the six units is shown in two ways. Above the diagonal are similarity indices (SI) for all possible comparisons, calculated as described¹⁰, measured in standard deviations above the mean (σ) for random matches and including gap penalties. Below the diagonal, the peptides are compared in terms of minimum base changes per codon (MBC) necessary to interconvert sequences¹⁰.

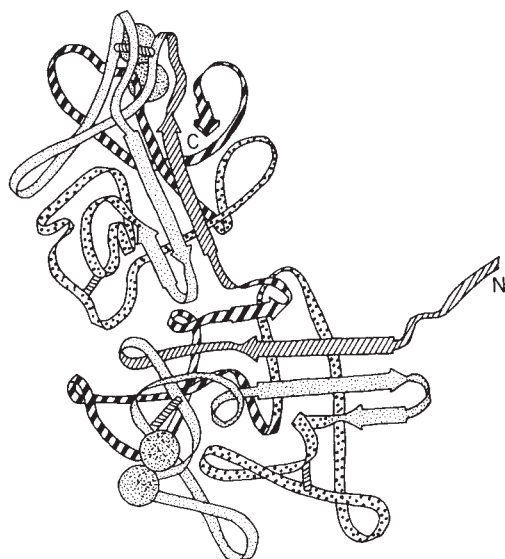
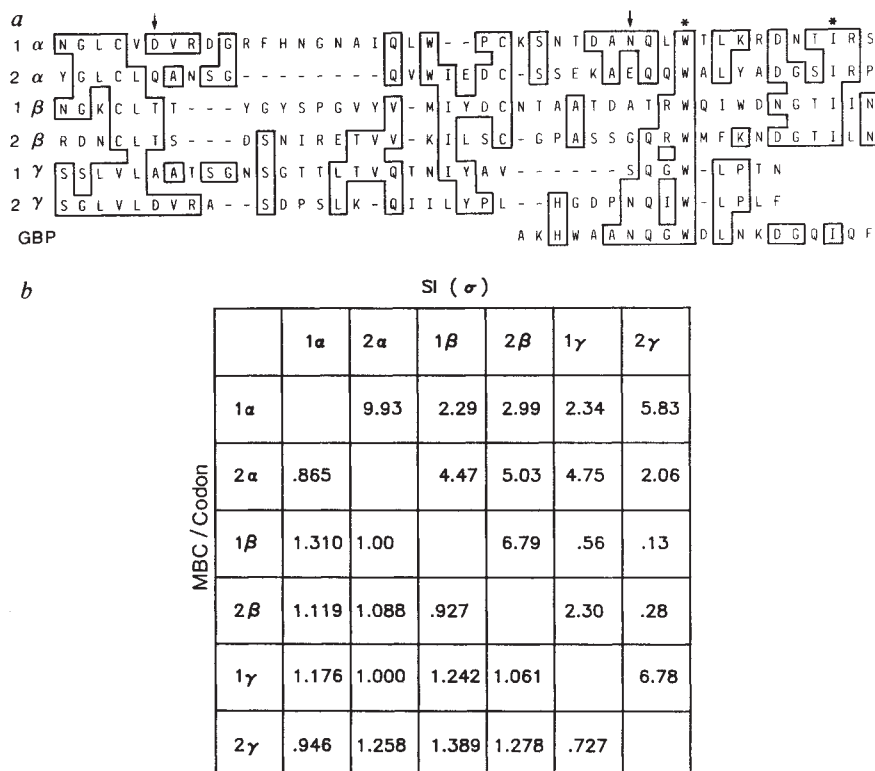


Fig. 2 Ribbon drawing of the ricin B chain. The B chain is viewed down the crystallographic *a* axis (and see ref. 15). The α -units have a small dot pattern, the β -units a large dot pattern and the γ -units have broad stripes. The λ -peptides are shown with narrow stripes. The positions of bound lactose sugars are shown as pairs of speckled circles.

subdomain can be described as a chain forming a loop, which binds galactose in the 1 α and 2 γ units, followed by a broad twist, and finishing with a second loop (or hook) structure.

The mode of galactose binding is nearly identical at the two sites. In both 1 α and 2 γ , lactose lies in a pocket formed on one side by a kink in the chain and on the other by an aromatic side-chain. The sequence of the kink, Asp-Val-Arg, is conserved in both subdomains. The aromatic groups are Trp 37 in 1 α and Tyr 248 in 2 γ . These side chains serve as a flat binding surface

for galactose but do not make more specific interaction with the sugar¹⁵. Galactose is hydrogen-bonded to homologous amides, Asn 46 and Asn 255, which are in turn stabilized by hydrogen bonds to homologous acids, Asp 22 and Asp 234. These binding residues are conserved only in those subdomains which still bind sugar, as indicated by the arrows in Fig. 1.

Other highly conserved residues can be assigned structural roles. The sequence Gln-X-Trp, where X is any residue, is found in five of the subdomains; the Trp itself is invariant. These Trp residues stabilize the carboxy-terminal hook of each subdomain by interacting either with highly conserved Ile residues downstream of the Trp (in the case of the α - and β -units) or with Ile residues from the λ -peptide (in the case of the γ -subdomains). These Trp-Ile van der Waals' contacts also stabilize the subdomain contacts within each domain, arranging the subdomains around a pseudo 3-fold axis (Fig. 4).

We hypothesize that the ancestor of the modern B-chain molecule was a galactose-binding peptide of ~40 residues which resembled the 1 α subdomain, the only unit which now retains all key structural and functional features. This primordial molecule probably existed as a trimer in solution, linked by the Trp-Ile interaction. Gene duplication and fusion then produced an $\alpha\beta\gamma$ -molecule similar to a modern B-chain domain. This fusion permitted specialization of the β -subdomains for a solely structural role; the $\alpha\beta\gamma$ -molecule thus probably bound two sugars at the α - and γ -positions.

Addition of the λ -peptide further stabilized the three fold structure. Another duplication then produced the current ($\lambda\alpha\beta\gamma$)₂ motif. This duplication blocked galactose access to the 1 γ unit, which retains only Gln 121 as a vestige of the sugar-binding site. Subdomain 2 α is accessible to solvent, but has also lost its binding site except for Gln 161. Thus the modern B chain binds two galactose moieties at sites separated by 75 Å on the extreme ends of the protein. Presumably selection for the two-domain structure involved proper spacing of the binding sites for cell surface binding and the triggering of endocytosis.

Our hypothetical primitive 'galactoside fold' is consistent in size both with the average exon-coded fragment seen in eukary-

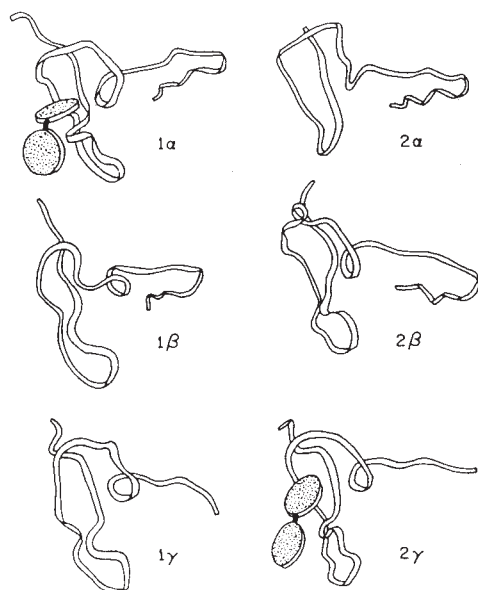


Fig. 3 Conformations of the individual subdomain units. The six subdomain units, presented as backbone ribbon drawings, have been positioned in the same orientation for comparative purposes.

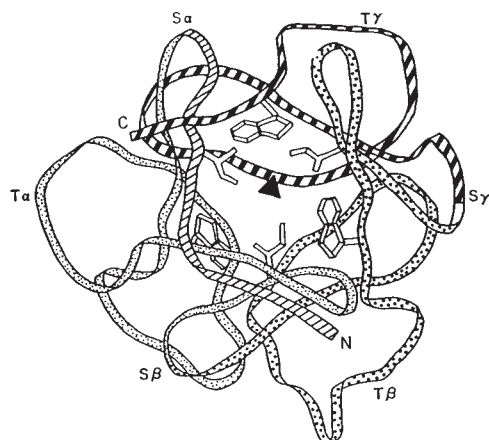


Fig. 4 The pseudo 3-fold arrangement of subdomain units in the amino-terminal domain of ricin B chain. The backbone of domain 1, plus the conservative Trp and Ile side chains are shown. Patterns to distinguish subdomains are as in Fig. 2. S, the start of each unit; T, the twist loop. The same folding pattern can be observed in domain 2 (not shown).

otic genomes, as suggested earlier¹⁰, and with the typical ligand-binding modules suggested to form the building blocks for assembly of protein domains¹¹. Interestingly, another plant lectin of known structure, wheat-germ agglutinin, is the product of repeated duplication and fusion of a different, similarly sized peptide¹². Stringing together multiple copies of a small sugar-binding fold would appear to be a simple and effective method for lectin formation.

It is possible, however, that our primitive 40-residue peptide was itself a duplication product. Traut has suggested that the simplest peptide folding units should be monofunctional¹¹, but our progenitor peptide is clearly bifunctional, binding both galactose and itself. The α -subdomains contain the repeated motif of a cysteine followed several residues later by the sequence Gln-X-Trp. Both halves have a simple loop structure. A peptide strikingly similar in sequence to a half subdomain can be seen in the structure of *Escherichia coli* galactose-binding

protein (GBP)¹³ (Fig. 1a), although this peptide has no apparent structural similarity to ricin subdomains and is not thought to be part of the binding site of the *E. coli* enzyme. Nevertheless, the idea of a simple short loop structure, capable of binding a biologically important ligand and acting as one kind of primordial molecular building block, is appealing. It is hard to imagine a stable biological element much smaller or simpler than such a loop; in the B chain we see how such loops can be strung together, self-assemble and be perturbed to create a modern functional protein of some complexity.

This research was supported by a grant from the NIH and by a gift from the Cetus Corporation of Emeryville, California. We thank Vivian Benningfield and Terrie Kolvoord for their help with the manuscript and Fred Hoffman for assistance with the figures.

Received 29 October 1986; accepted 9 February 1987.

1. Olsnes, A. & Pihl, A. in *The Molecular Actions of Toxins and Viruses* (eds Cohen, P. & Van Heynigen, S.) 52-105 (Elsevier Biomedical, New York, 1982).
2. Nicolson, G. L. & Blaustein, J. *Biochim. biophys. Acta* **266**, 543-547 (1972).
3. Baenziger, J. U. & Fiete, D. *J. biol. Chem.* **254**, 9795-9799 (1979).
4. Nicolson, G. L., Lacorbriere, M. & Hunter, T. R. *Cancer Res.* **35**, 144-155 (1975).
5. Zentz, C., Frenoy, J. P. & Bourrillon, R. *Biochim. biophys. Acta* **536**, 18-26 (1978).
6. Houston, L. L. & Dooley, T. P. *J. biol. Chem.* **257**, 4147-4151 (1982).
7. Mise, T., Shimoda, T. & Funatsu, G. *Agric. Biol. Chem.* **50**, 151-155 (1986).
8. Hatakeyama, T., Yamasaki, N. & Funatsu, G. *J. Biochem.* **99**, 1049-1056 (1986).
9. Villafranca, J. E. & Robertus, J. D. *J. biol. Chem.* **256**, 554-556 (1981).
10. Robertus, J. D. & Ready, M. P. *J. biol. Chem.* **259**, 13953-13956 (1984).
11. Traut, T. W. *Mol. cell. Biochem.* **70**, 3-10 (1986).
12. Wright, H. T., Brooks, D. M. & Wright, C. S. *J. molec. Evol.* **21**, 133-138 (1985).
13. Mahoney, W. C., Hogg, R. W. & Hermodson, M. A. *J. biol. Chem.* **256**, 4350-4356 (1981).
14. Halling, K. C. *et al. Nucleic Acids Res.* **13**, 8019-8033 (1985).
15. Montfort, W. *et al. J. biol. Chem.* (in the press).

Corrigendum

Lead concentration changes in Antarctic ice during the Wisconsin/Holocene transition

Claude F. Boutron & Clair C. Patterson
Nature **323**, 222-225 (1986).

THERE are two misprints in the published version of this paper: Page 223, in line 26 of the section 'Character of the data', the age of the Wisconsin blue ice block is >10,000 years, not 10,000 years. Line 27 of the same section: lead concentration in this block is 1.75 pg Pb g⁻¹, not <1.75 pg Pb g⁻¹.

There is also some ambiguity in the way that depths are referred to. In most parts of the paper (including Table 1), depths are expressed as *real depths*, but in lines 9-11 of the 'Experimental techniques' section, depth limits for the four climatic stages taken from ref. 27 were expressed as metres of ice equivalent. Transforming these depths into *real depths* the depths of the limits of the four climatic stages become:

- (1) Beginning of the Holocene stage: ~408 m (real depth) instead of ~380 m (ice equivalent depth given in Lorius *et al.*²⁷).
- (2) Beginning of the Wisconsin/Holocene transition: ~548 m (real depth) instead of ~520 m (ice equivalent).
- (3) Beginning of the Last Glacial Maximum: ~708 m (real depth) instead of ~680 m (ice equivalent). The corresponding paragraph then reads:

"The oxygen isotope profile of Lorius *et al.*²⁷ shows four main climatic stages: the Holocene stage from the surface down to ~408 m, the Wisconsin/Holocene transition from ~408 to 548 m, the Last Glacial Maximum from ~548 to 708 m, and the earlier Wisconsin stage from ~708 m to the bottom of the core".

Finally, the vertical bars at the top of Fig. 2, p. 224, which show the approximate limits of the four climatic stages are slightly erroneous: they should be slightly shifted toward the right: the Holocene/Transition limit should be at about 10,900 years BP instead of about 10,000 BP; the Transition/Wisconsin limit should be at about 15,800 yr BP instead of about 15,000 yr BP. □